

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS
Departamento de Física Atómica, Molecular y Nuclear



TESIS DOCTORAL

Sensitivity studies for the Cherenkov Telescope Array

Estudios de sensibilidad para el Cherenkov Telescope Array

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Tarek Hassan Collado

Advisors

José Luis Contreras González
Néstor Mirabal Barrios

Madrid, 2016

SENSITIVITY STUDIES FOR THE CHERENKOV TELESCOPE ARRAY

Estudios de sensibilidad para el Cherenkov Telescope Array



Tarek Hassan Collado

Departamento de Física Atómica, Molecular y Nuclear
Universidad Complutense de Madrid

Thesis advisors

Dr. José Luis Contreras González

Dr. Nestor Mirabal Barrios

A thesis submitted for the degree of

Doctor of Philosophy

June 2015, Madrid

Para Papá.

Acknowledgements

I would like to express my sincere gratitude to every single person I got involved with during this 4 years of PhD. I enjoyed it as much as I did thanks to you all.

First and foremost, I want to thank Jose Luis and Nestor, for giving me this live-changing opportunity. I hope someday you will feel as proud to have me as a student as I am for having you as my advisors. This work, and those to come, would be impossible without you.

All members of the *Grupo de Altas Energías* helped immensely to this work. Victoria, Konstancja, Valeria, José Luis, Alejandro, Juan Abel, Emilio and Marcos; to be surrounded by you all was not only a great help to my work, but also a motivation in the PhD pursuit. I specially want to thank my colleagues and dear friends: Daniel, only now I understand how thankful I should be for your help during your last year of PhD. I also thank you for my time in New York, as it was one of the three best things of my doctorate. Luis Ángel and Diego, as my consultants of crazy ideas, I will definitely miss my constant visits to the lab. I cannot think about better colleagues to have shared these years with as Simon and David, great scientists, coworkers and friends (despite their nationalities). I also need to thank Ester, as she was the (only) positive side of bureaucracy, and Miguel, who will finally be able to rest without my questions and nagging.

Also, I need to acknowledge the people I had the pleasure to meet and work with: In regards to the Columbia University Physics Department, I need to thank Manel and Marcos for the warm welcome and support during my stay, and specially to Brian, for making the amazing experience possible. I am also grateful to everyone at IFAE, particularly

to Abelardo, as he was a great help and his contribution to this work is essential. Surprisingly, I am also thankful to Alba, Carlos, Carlitos, Dani, Eudald, John E., Leyre, Paolo, Pauline, Quim, Raúl and Rubén, who (unfortunately) treated me as one of their own.

Finally, I would like to thank my family: Sam, Rana, Mom and Dad for your never-ending support, constant cheers and care. Also, to my other family from *el jardín* who suffered my company during the tough times. Lastly, thank you Eva for your support, encouragement, patience, love, and everything you give me every single day. I owe you far more than this work. Thank you.

Contents

Contents	i
Introduction	1
1 High Energy Astrophysics	5
1.1 Cosmic rays	6
1.2 Production and absorption processes	8
1.2.1 Absorption mechanisms	12
1.2.2 WIMP annihilation	14
1.3 Gamma ray astronomy	16
1.3.1 Known gamma ray emitters	16
1.3.1.1 Galactic γ -ray emitters	16
1.3.1.2 Extragalactic γ -ray sources	20
1.3.2 Potential WIMP annihilating targets	24
2 Past, present & future of γ-ray detectors	27
2.1 Historical motivations	28
2.2 Observation Techniques	29
2.2.1 Space detectors	31
2.2.2 Ground detectors	32
2.2.3 Ground vs Space	34
2.3 Current γ -ray observatories	37
2.3.1 Fermi Gamma-ray Space Telescope	37
2.3.1.1 Instruments on board	39
2.3.1.2 Main results	40

CONTENTS

2.3.2	Imaging Atmospheric Cherenkov Telescopes	43
2.3.2.1	Scientific results	45
2.4	Future γ -ray observatories	47
2.4.1	HAWC	48
2.4.2	CTA	49
2.4.2.1	CTA Key Science Projects	51
2.4.2.2	CTA performance	53
3	Sensitivity studies for the CTA	57
3.1	Imaging Atmospheric Cherenkov Telescopes	58
3.1.1	Technical description	58
3.1.2	Stereoscopic imaging principle	62
3.1.3	CTA telescopes	64
3.2	Monte Carlo simulation tools	67
3.2.1	CORSIKA	70
3.2.2	<code>sim_telarray</code>	71
3.3	CTA Analysis tools	74
3.3.1	Data analysis overview	74
3.3.2	Alternative CTA MC analysis methods	77
3.3.3	CTA MARS based analysis	79
3.3.3.1	Trace integration	80
3.3.3.2	Image cleaning and parameterization	81
3.3.3.3	Stereo reconstruction	84
3.3.3.4	Performance estimation	88
3.4	CTA Monte Carlo productions	91
3.4.1	Prod-1	91
3.4.2	Prod-2	95
3.4.2.1	Layout candidates	100
3.4.2.2	Layouts with mixed MST types	102
3.5	CTA performance	102
3.5.1	Plausible candidates performance	105
3.5.2	IRF optimization	111
3.5.3	Layout design study of mixed MST types	113

3.5.3.1	SC-MST telescope performance	113
3.5.3.2	Telescope spacing and distribution	117
3.5.3.3	Off-axis performance	121
3.6	CTA-N site selection	122
3.6.1	Site comparison	124
3.6.2	Averaged Geo-magnetic field	126
3.6.3	Tenerife NSB	130
3.6.4	La Palma alternative layouts	132
4	CTA forecast	139
4.1	Physics evaluation tools	140
4.1.1	<i>GAETools</i>	141
4.1.2	Comparison with <i>CTA macros</i>	143
4.1.3	<i>Fermi</i> catalog extrapolations	144
4.1.4	DAMASCO	146
4.2	Physics cases	147
4.2.1	Galactic sources and surveys	148
4.2.2	Extragalactic follow-up observations	151
4.2.2.1	CTA-N impact	155
4.2.3	Pulsars	157
4.2.4	Prospects for Dark Matter searches in dSph galaxies	162
5	Applications of Machine Learning Algorithms to the High Energy Astrophysics	167
5.1	Introduction to Machine Learning Algorithms	167
5.1.1	Random Forest	169
5.1.2	Support Vector Machines	170
5.2	MVA in CTA analysis	170
5.2.1	Energy reconstruction	171
5.2.2	Gamma-Hadron separation	173
5.3	MVA for source type determination in 2FGL	177
5.3.1	Sybil	178
5.3.1.1	Datasets	179

CONTENTS

5.3.1.2	Application to unassociated sources	181
5.3.1.3	Search for dark matter subhaloes in the 2FGL . . .	182
5.3.1.4	Discussion	185
5.3.2	AGN type determination	186
5.3.2.1	Datasets	187
5.3.2.2	Results	190
5.3.2.3	Outlier detection and potential biases	193
5.3.2.4	Application to unassociated <i>Fermi</i> objects	194
5.3.2.5	Discussion	196
6	Conclusions	199
A	Extended Atmospheric Showers	205
A.1	The Cherenkov radiation	205
A.2	Extended atmospheric showers	208
A.2.1	Gamma-ray induced extended atmospheric showers	209
A.2.2	Hadron induced extended atmospheric showers	212
A.2.3	Electron induced extended atmospheric showers	215
A.3	Geomagnetic field effect on <i>EAS</i>	217
B	Sensitivity studies for the Cherenkov Telescope Array	221
C	Estudios de sensibilidad para el Cherenkov Telescope Array	223
	List of Figures	225
	List of Tables	231
	References	233

Introduction

Since the creation of the first telescope in the 17th century, every major discovery in astrophysics has been the direct consequence of the development of novel observation techniques, opening new windows in the electromagnetic spectrum. After Karl Jansky discovered serendipitously the first radio source in 1933, Grote Reber built the first parabolic radio telescope in his backyard, planting the seed of a whole new field in astronomy. Similarly, new technologies in the 1950s allowed the establishment of other fields, such as the infrared, ultraviolet or the X-rays.

The highest energy end of the electromagnetic spectrum, the γ -ray range, represents the last unexplored window for astronomers and should reveal the most extreme phenomena that take place in the Universe. Given the technical complexity of γ -ray detection and the extremely relative low fluxes, γ -ray astronomy has undergone a slower development compared to other wavelengths. Nowadays, the great success of consecutive space missions together with the development and refinement of new detection techniques from the ground, has allowed outstanding scientific results and has brought gamma-ray astronomy to a worthy level in par with other astronomy fields.

This work is devoted to the study and improvement of the future Cherenkov Telescope Array (CTA), the next generation of ground based γ -ray detectors, designed to observe photons with the highest energies ever observed from cosmic sources.

The thesis is arranged as follows:

- **Chapter 1:** An introduction to the high-energy astrophysics field. It describes the different γ -ray production and absorption mechanisms, together with known and potential astronomical sources expected to emit in this en-

ergy range.

- **Chapter 2:** A summary of the history of γ -ray detectors and the different detection techniques employed to measure highly energetic photons. Both ground based and space detection techniques are described, comparing their performance and summarizing their main results. This chapter concludes describing future γ -ray experiments, describing their capabilities and their scientific potential.
- **Chapter 3:** This chapter is devoted to sensitivity studies performed for the CTA. I describe the Imaging Atmospheric Cherenkov Telescope technique, followed by an explanation of the Monte Carlo (MC) simulation tools used. The peculiarities of the analysis used in this work are presented and compared with alternative tools developed by the CTA collaboration. I also discuss the implications of our results for telescope distribution, telescope types and possible improvements in future data analysis. Finally, results concerning the selection of the CTA construction site are shown, studying the effect of the construction site on performance.
- **Chapter 4:** This chapter considers realistic physics cases to assess and evaluate CTA capabilities, in order to gauge the effect of performance differences over real scientific output. An introduction to the software developed in this work is given, comparing attained results with other tools developed by the collaboration. Then, forecasts are performed for several scientific topics such as source populations, blazars and pulsars detectability or Dark Matter prospects. The performance over these topics of different telescope layout candidates and construction site altitudes is also evaluated.
- **Chapter 5:** This last chapter describes additional applications of machine learning techniques. A short introduction to the algorithms applied is given, describing their different uses within the CTA analysis and improvements over alternative methods. Then, results on γ -ray source type determination are presented, using spectral features in the very high energy range to determine if cosmic objects of uncertain type are pulsars or blazars, sub-categorizing the blazar class into flat-spectrum radio quasars (FSRQ) or BL

Lacertae (BL Lac).

- **Chapter 6:** This chapter provides a summary of the main conclusions of this work as a whole. Results are discussed with a particular emphasis on open improvements to be performed in the future.
- **Appendix A:** This appendix introduces the theory behind the expended air showers generated by impinging γ -rays. First, it describes the Cherenkov radiation taking place in the Earth's atmosphere. Then, the different types of generated showers are described, regarding the primary particle initiating them. Although it may not be considered essential, the reader is encouraged to read it, as an understanding of these showers is required to comprehend site performance differences with respect to altitude and the geo-magnetic field.
- **Appendix B:** This appendix provides the required summary of this work in the English language.
- **Appendix C:** The last appendix provides the required summary of this work in the Spanish language.

Chapter 1

High Energy Astrophysics

After the first half of the XX century, a whole new universe was opened to scientists all over the world expanding the observable universe to other wavelengths of the electromagnetic spectrum. Akin to the radio, infrared, ultraviolet and X-ray bands, the γ -ray range represents the last window to explore. With the detection of exotic objects such as quasars, pulsars, blazars, radio galaxies, neutron stars or black holes, a stream of new theories emerged to explain their existence and observed properties.

These theories anticipated the existence of extremely energetic processes in the universe, such as supernova explosions or cosmic ray interactions with interstellar gas, capable of emitting photons carrying huge amounts of energy. The γ -ray astronomy was born after the spacecraft technology boost consequent to the Second World War, deploying the first γ -ray detectors into orbit. Unlike the other wavelengths, the universe these detectors observed could not be explained with conventional thermal emission. The high energy astronomy field studies these *non-thermal* processes occurring in the universe.

As described in chapter 2, the high energy astronomy grew hand in hand with the particle physics, studying the cosmic rays nature, due to the similar methods applied to detect them. This is the reason why the field is generally referred as *astroparticle physics*. Nowadays, after the development of the current observation techniques from ground and space, *astroparticle physics* is becoming a worthy new field in astronomy while offering an exceptional area of research in cosmology and

fundamental physics. Within this field, the emission models and nature of the most extreme astronomical objects is studied, together with topics related to the fundamental physics such as the detection of Dark Matter, the search of axion-like particles or constraining the extragalactic background light models.

This chapter will give a shallow overview of some topics studied within the field of astroparticle physics. For further knowledge of any of these topics, the reader is encouraged to follow the references and the following bibliography [1, 2, 3]. First, in section 1.1, our current knowledge of the cosmic-rays is summarized and their current believed origin. Then, in section 1.2, an introduction is given to the different processes capable of emitting and absorbing γ -rays, crucial to understand this field. Lastly, in section 1.3, the main sources of γ -ray astronomy are outlined, together with some potential targets still undetected.

1.1 Cosmic rays

During the first half of the XX century, this term was used for referring to both the highly energetic electromagnetic radiation and the nuclei of cosmic origin. Nowadays, the "cosmic rays" term, defined after Robert Millikan [4] improved V. Hess [5] measurements, refers to the high energy charged particles originated in outer space which constantly collide with the Earth's atmosphere. These particles have been deeply studied for more than a century, in order to explain their origin, still under debate, and the processes that could accelerate these particles to such enormous energies, far exceeding those reached by the current generation of particle accelerators.

The composition of the cosmic-ray spectrum, as shown in Fig. A.6, depends on the energy considered, but is composed essentially by protons (79%) and helium (15%). The remaining 6% corresponds to high energy electrons and heavier nuclei. The ensemble of all these particles as a function of the energy forms the cosmic-ray spectrum, shown in Fig. 1.1, extending through more than 12 orders of magnitude. The green dashed line represents the usual fit used to describe it, with an slope of -2.7.

As highlighted in Fig. 1.1, the spectrum shows 2 inflection points: the *knee*, located at $\approx 4 \times 10^{15}$ eV, softening the spectrum and the *ankle*, at $\sim 10^{18}$ eV,

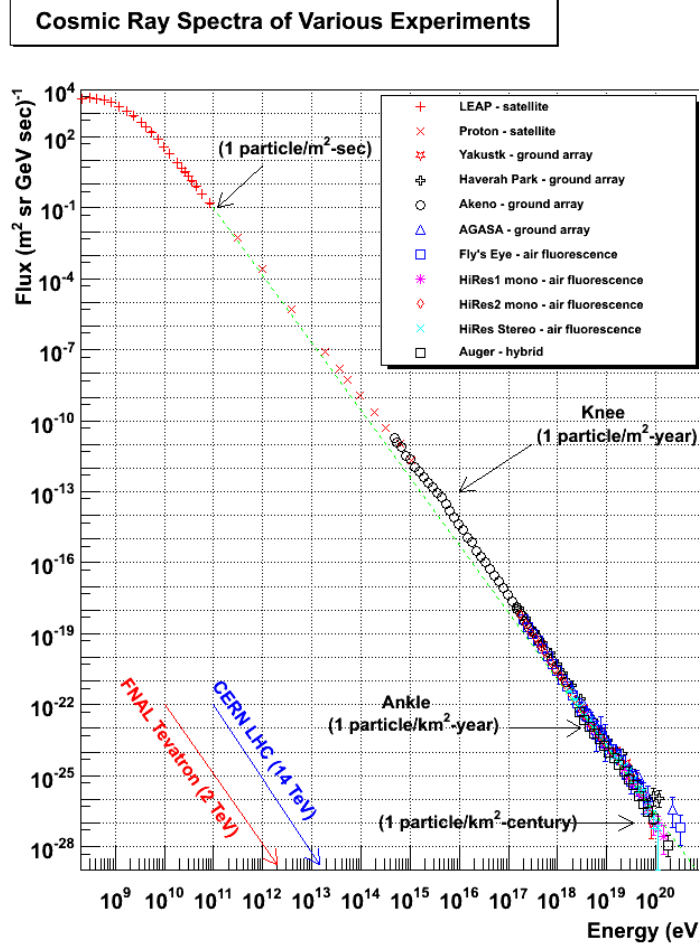


Figure 1.1: Cosmic-ray flux as a function of the particle's energy. Courtesy of Dr. William Hanlon [6].

hardening it. In the highest energy end of the spectrum, at $\approx 4 \times 10^{19} \text{ eV}$, the interaction with the Cosmic Microwave Background (CMB) radiation produces the Greisen-Kuzmin-Zatsepin (GKZ) cutoff. This distribution is currently explained separating the cosmic ray spectrum into two different contributions: the galactic and extragalactic cosmic rays.

- Galactic cosmic-rays: Below 10^9 eV , the spectrum is well described with Solar Wind emitted by the Sun, showing chemical abundances consistent with the ones observed in the star. Particles with energies up to $\sim 10^{18} \text{ eV}$ are generally ascribed to the cosmic rays accelerated within our galaxy. The

1. HIGH ENERGY ASTROPHYSICS

most clear candidate for galactic cosmic-ray accelerators are the Supernova Remnants (SNRs), expected to accelerate protons up to $\sim 10^{15}$ eV (which could explain the observed decreased flux after the *knee*), although other sources like micro-quasars and pulsars were also proposed.

- Extragalactic cosmic-rays: The *ankle* is attributed to the point where the extragalactic component of the cosmic ray spectrum becomes predominant. At these energies, particles are no longer confined by galactic magnetic fields, which could explain the spectral hardening after the *ankle*. Several candidates have been proposed to produce particle acceleration above $\sim 10^{18}$ eV: Active Galactic Nuclei (AGN), γ -ray bursts, radio galaxy lobes, or magnetic fields in galaxy clusters. Charged heavy particles with such extreme energies are not greatly affected by magnetic fields, and a rough direction estimation could be reconstructed, carrying information of the region where they were accelerated. This kind of approach was used by the Pierre Auger experiment (see 2.2.2) to correlate incoming cosmic rays with nearby AGNs [7, 8].

1.2 Production and absorption processes

To comprehend the non-thermal universe is essential to have a deep understanding of the different processes capable of producing gamma ray emission. Here a short outlook of the most relevant γ -ray production mechanisms is given. These processes are the following (shown in Fig. 1.2):

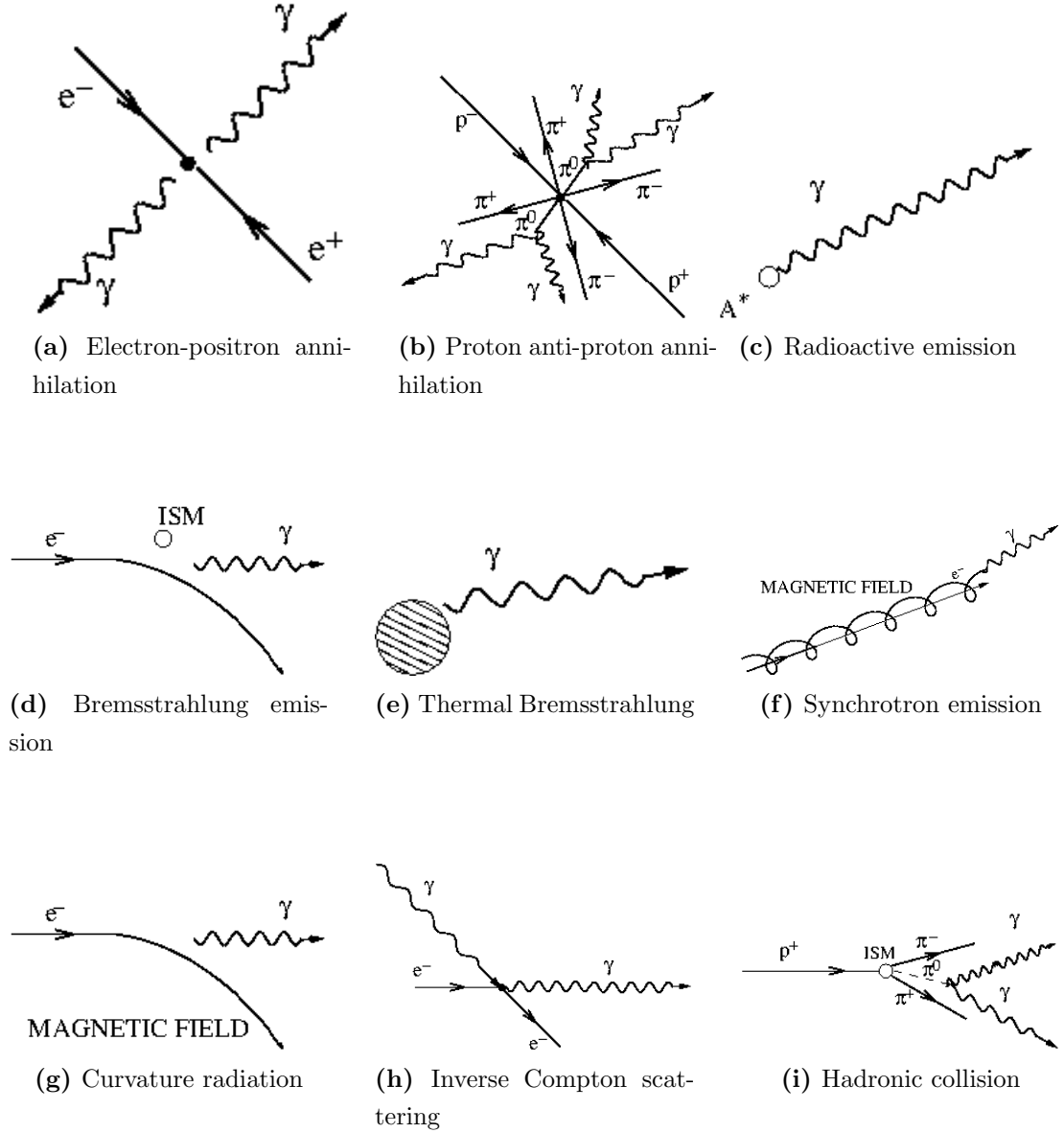


Figure 1.2: Schematic diagrams of γ -ray emission processes. Simplified diagrams of γ -ray production mechanisms. Figures extracted from [9].

- **Matter anti-matter annihilation:** Particle annihilation through matter anti-matter pairs processes, mainly e^+e^- , are able to create γ -rays. The electron positron ($e^+e^- \rightarrow 2\gamma$) (1.2a) directly produce γ -ray emission, with an energy equal to the mass of the particles involved in the process. Unless their kinetic energy is comparable to their rest energy, γ -rays produced are

1. HIGH ENERGY ASTROPHYSICS

typically $E_\gamma = m_e c^2 = 511 \text{ keV}$. Proton anti-proton processes (1.2b) are also expected to occur indirectly producing γ -rays, through neutral pions decay, although the contribution of this process is considered rather marginal.

- **Radioactive nuclei:** Together with the e^+e^- annihilation, radioactive emission (1.2c) dominate the low energy γ -ray ($\sim \text{MeV}$) production. Both radioactive decay or collisional excitation processes are able to produce high energy photons.
- **Bremsstrahlung emission:** Bremsstrahlung emission from cosmic ultra-relativistic electrons is the dominant process for galactic diffuse γ -ray production, up to $\sim 100 \text{ MeV}$. The γ -ray emission is produced as a charged particle is deflected in the vicinity of a nucleus (or ionized atom) due to its electromagnetic field. As this process is proportional to the ratio between the particle mass and charge, electrons are far more efficient than protons, dominating γ -ray production through this kind of emission. Hot ionized gasses (1.2e) produce this kind of processes with energies following Wien's law ($E_{max} (\text{MeV}) = 4.7 \times 10^{-10} T (K)$), requiring temperatures of the order of 10^{10} K to have a significant contribution of γ -rays above 1 MeV . Such temperatures cannot be stable (they are 2 orders of magnitude above hydrogen and helium fusion) and are considered unusual, making thermal Bremsstrahlung contribution to the γ -rays not very relevant.
- **Synchrotron emission:** Together with curvature radiation, synchrotron emission (1.2f) takes place when ultra-relativistic electrons move immersed in a magnetic field. These electrons describe helicoidal trajectories following the magnetic field lines. Synchrotron photons are emitted through an angle $\theta \approx m_e c^2 / E$ pointed towards the electron circular motion. The maximum emission will take place at energies:

$$E_{\gamma, max} (\text{eV}) = 5 \times 10^{-9} H_\perp [E_e / (m_e c^2)]^2, \quad (1.1)$$

where H_\perp is the perpendicular component of the magnetic field to the electron movement and E_e its energy. Unrealistic high values of E_e and H are

required to produce high energy γ -rays, but ultra-relativistic electrons are able to emit photons with lower energies and then acquire much higher energies through inverse Compton interaction, called Synchrotron self-Compton (SSC).

- **Curvature radiation:** If charged particles travel through magnetic field lines with curved trajectories and huge intensity (10^{11} - 10^{13} G), photons are emitted parallel to the magnetic field lines (1.2g). Unlike in the synchrotron emission, the radius producing the rotation is ρ_c , the curvature radius of the magnetic field lines. Emitted photons have energies of the order of:

$$E_\gamma \text{ (eV)} \approx (3/2)\hbar c\gamma^3/\rho_c = 2.96 \times 10^{-5}\gamma^3/\rho_c \text{ (cm)}, \quad (1.2)$$

where γ is the Lorentz factor of the electron. This emission mechanism is relevant in environments with extreme magnetic field intensities, such as in the vicinity of pulsars, producing γ -rays with energies reaching the GeV.

- **Inverse Compton:** This term refers to the Compton interaction in which the photon gains energy at the expense of an ultra-relativistic electron. This process is highly efficient, and is able to produce high energy γ -rays from photons travelling through a dense population of relativistic electrons. If a photon travels with an energy $E_\gamma^0 \ll m_e c^2$, then the centre of momentum frame is very close to the rest frame of the relativistic electron. The energy of the generated photon will be Doppler shifted:

$$E_\gamma = \gamma E_\gamma^0 [1 + (v/c) \cos \theta], \quad (1.3)$$

where γ is the Lorentz factor, v is the velocity of the electron and θ is the angle of incidence. Now considering the interaction between Cosmic Microwave Background (CMB) photons ($E_\gamma^0 \approx 6 \times 10^{-4}$ eV) with a population of relativistic photons, the average energy of the scattered photons is $\langle E_\gamma \rangle \approx 4/3\gamma^2 E_\gamma^0$ with a maximum energy of $E_\gamma^{max} \approx 4\gamma^2 E_\gamma^0$. Considering the population of relativistic electrons follow a power law distribution $dN_e/dE \propto E^{-\alpha}$, the spectral shape of the resulting boosted photons will

1. HIGH ENERGY ASTROPHYSICS

also follow the same distribution: $dN_\gamma/dE \propto E^{-(\alpha+1)/2}$

- **Hadronic collisions:** Multitude of processes, such as strong magnetic fields or jet structures, are able to accelerate particles and produce very energetic collisions. Those which are hadronic, mostly between accelerated protons and heavier nuclei (1.2i), are able to generate many kinds of secondary particles, mainly pions (π^0 , π^+ , π^-), and in a lower degree kaons (K^+ , K^-) and other nucleons. Some of these particles decay in high energy γ -rays. The most efficient γ -ray emitter in these processes is the neutral pion π^0 , with a probability of 99.8% to annihilate as $\pi^0 \rightarrow \gamma\gamma$. Proton-proton collisions usually produce a single neutral pion (as long as they hold enough energy to produce them: $E_{th} = 2m_{\pi^0}c^2(1 + m_{\pi^0}/4m_p) \simeq 280 \text{ MeV}$) emitting γ -ray spectral distributions similar to the parent proton population. Cosmic rays interacting with the interstellar medium are responsible of most of the galactic diffuse γ -ray radiation above 100 MeV. Also, the annihilation of neutral pions is the dominant process in some γ -ray sources, such as supernova remnants or in Active Galactic Nucleus (AGN)s according to some models.

1.2.1 Absorption mechanisms

In addition to the processes participating in the γ -ray production, the mechanisms involved in the absorption of these high energy photons is of crucial importance to astronomers, in order to infer the intrinsic γ -ray flux emitted by sources from the measured spectrum. The process governing γ -ray absorption is the **pair production**.

Photon-photon interactions are able to produce an electron-positron pair if the global energy is larger than the rest energy of both particles $E_{\gamma\gamma} > 2m_e c^2$. There are two possible scenarios where γ -rays can produce e^-e^+ pairs:

- **Classical pair production:** A single photon traveling through the vicinity of a charged particle is able to produce a e^-e^+ pair if it holds enough energy ($2m_e c^2$) via annihilation with a virtual photon from the particle's field: $\gamma\gamma_{\text{virtual}} \rightarrow e^+e^-$. The classical pair production cross section for energies

larger than 30 MeV is the following (extracted from [10]):

$$\sigma_{\gamma\gamma} = \sigma_0 Z^2 \left[\frac{28}{9} \ln \left(\frac{183}{Z^{1/3}} \right) - \frac{2}{27} \right], \quad (1.4)$$

where $\sigma_0 = (1/137)(e^4/m_e c^4)$ and Z is the atomic number. This process is responsible of the γ -ray absorption in the atmosphere (generating an Extended Air Shower (EAS), see appendix A) and is exploited by γ -ray detectors above the Earth's atmosphere (see 2.2.1).

- **γ - γ pair production:** In this case, 2 real photons with a total energy larger than $2m_e c^2$ collide and produce the $e^- e^+$ pair: $\gamma_{\text{VHE}} \gamma_{\text{soft}} \rightarrow e^- e^+$. In this case, the cross section of a γ -ray photon of energy E_γ with a background photon of energy ϵ is (extracted from [11]):

$$\sigma_{\gamma\gamma} = \frac{3\sigma_T}{16}(1 - \beta^2) \left[2\beta(\beta^2 - 2) + (3 - \beta^4) \ln \left(\frac{1 + \beta}{1 - \beta} \right) \right], \quad (1.5)$$

$$\beta = \left(1 - \frac{\epsilon_{th}}{\epsilon} \right)^{1/2}, \quad \epsilon_{th}(E_\gamma, \theta) = \frac{2(m_e c^2)^2}{E_\gamma(1 - \cos\theta)}, \quad (1.6)$$

where $\sigma_T = 6.65 \times 10^{-25} \text{ cm}^2$ is the Thomson cross section, ϵ_{th} the threshold energy of the interaction and θ the angle between the incident photons.

Photon-photon pair production is of critical importance for astronomers in the Very High Energy (VHE) range. Low energy photons from the Extragalactic Background Light (EBL) collide with VHE γ -rays, producing an absorption in the observed spectrum, increasing with the distance (redshift). This attenuation can be described by an optical depth τ , as a function of the energy of the photon E_γ and the redshift z (see Fig. 1.3):

$$F(E) = F_0(E) e^{-\tau(E_\gamma, z)}, \quad (1.7)$$

where $F(E)$ is the measured flux and $F_0(E)$ is the intrinsic spectrum of the source. In one hand, as the EBL absorption has been modeled as a function of the distance [12, 13, 14], if the redshift of a source is known, the intrinsic

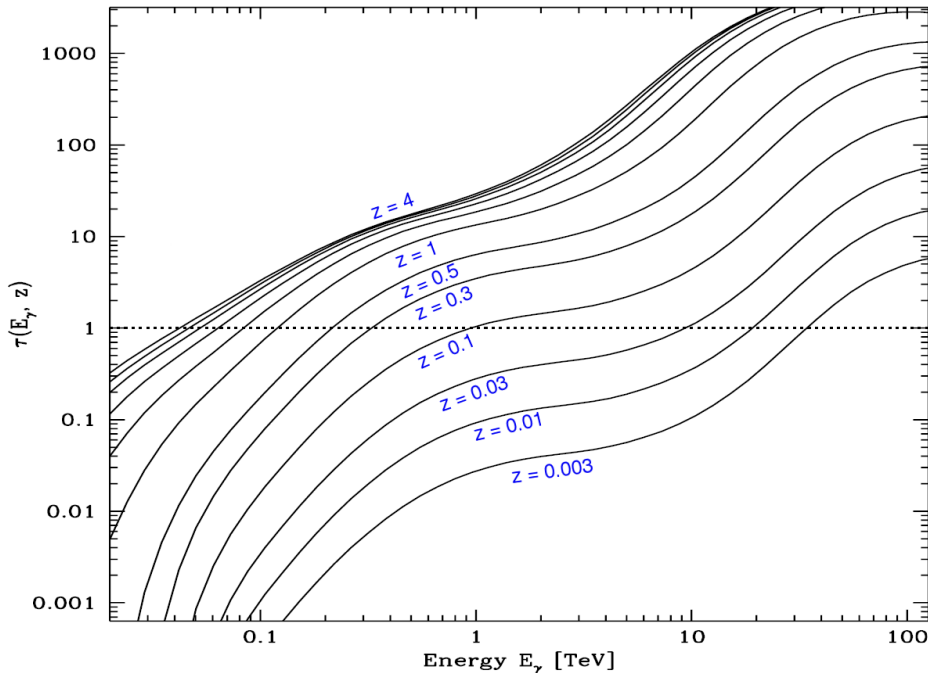


Figure 1.3: The optical depth by photon-photon pair production as a function of the photon energy E_γ for sources located at different redshifts: $z = 0.003, 0.01, 0.03, 0.1, 0.3, 0.5, 1, 1.5, 2, 2.5, 3, 4$. Modified version of Fig. 7 from [12].

spectrum can be estimated, as described in section 4.1.1. On the other hand, observations in the VHE range of distant sources are able to experimentally constraint and characterize the EBL [15], or even help to estimate the redshift of distant blazars.

1.2.2 WIMP annihilation

One of the most important questions still open in astronomy is the nature of the Dark Matter (DM). A large number of evidences, measuring cosmic objects motion in all kinds of distance scales, seem to point in the same direction: There is a very significant part of the universe of unknown nature that we have not detected yet, able to interact with ordinary matter through gravitational forces. In fact, recent observations performed by the Wilkinson Microwave Anisotropy Probe (WMAP) satellite [16] quantified the DM component to account for the 25% of the total

energy of the whole Universe.

Large scale galaxy formation simulations were performed to compare the observed large redshift galaxies with the replicated one [17, 18, 19, 20]. These simulations predicted the cold Dark Matter (CDM) scenario: DM particle candidates decoupled from the thermal equilibrium in the early universe (*freeze-out*) and moved with non-relativistic velocities. These conclusions and other few assumptions form the most popular scenario for CDM: DM is composed by a Weakly Interacting Massive Particle (WIMP), only affected by weak interactions, with a mass ranging between few tens of GeV up to few TeV [21, 22]. Currently, none of the known particles from the standard model have consistent properties with the ones required for WIMPs. Certain SuperSYmmetry (SUSY) theories predict feasible WIMP candidates that could annihilate into standard model particles as their final states.

Astroparticle physics may have something to contribute in this topic. Similarly as the e^-e^+ annihilation signature observed at 511 keV, an excess with a specific shape in the VHE range observed in sources of known high DM density could lead to the first ever detection of the constituents of DM. Following the terminology of [23], the expected γ -ray differential flux within a solid angle $\Delta\Omega$ from an astronomical DM rich target is:

$$\frac{d\Phi(\Delta\Omega, E_\gamma)}{dE_\gamma} = B_F \cdot \frac{1}{4\pi} \underbrace{\frac{(\sigma_{\text{ann}}v)}{2m_\chi^2} \sum_i \text{BR}_i \frac{dN_\gamma^i}{dE_\gamma}}_{\text{Particle Physics}} \cdot \underbrace{\tilde{J}(\Delta\Omega)}_{\text{Astrophysics}}, \quad (1.8)$$

where m_χ is the mass of the WIMP particle, $(\sigma_{\text{ann}}v)$ is the annihilation cross-section multiplied by the relative velocity of the two WIMPs, $\sum_i \text{BR}_i dN_\gamma^i/dE_\gamma = dN_\gamma/dE_\gamma$ is the sum over all annihilation channels of each individual photon flux i with branching ratios BR_i . The *astrophysical factor* \tilde{J} , is the integral of the square DM density along the line of sight (los) and solid angle $\Delta\Omega$. The *boost factor* B_F accounts for all the possible contributions that could increase the generated γ -ray flux, such as inhomogeneities in the DM profile density.

1.3 Gamma ray astronomy

Driven by the γ -ray satellite telescopes and the current generation of ground based detectors, the last 20 years have revolutionized the γ -ray astronomy. The number of detected sources in the high energy range has increased exponentially as a result of the great success of the Fermi Gamma-ray Space Telescope, identifying up to 3033 γ -ray point like sources [24]. From ground, observing the highest end of the electromagnetic spectrum, 156 detections have been achieved, although this number is expected to dramatically increase with the new generation of ground based detectors, High-Altitude Water Cherenkov (HAWC) and the CTA (see section 2.6).

In this section, a short description of the known and potential γ -ray emitters is given. For further knowledge on γ -ray sources and the main emission processes involved in each case, the reader is encouraged to read references such as [3], or related thesis dissertations cited in each topic.

1.3.1 Known gamma ray emitters

Two major classes of γ -ray emitters are described: galactic and extragalactic γ -ray emitters.

1.3.1.1 Galactic γ -ray emitters

Just with a quick look at the sky in the High Energy (HE) range (see Fig. 2.5), is clear that the richest region of γ -ray emitters is located in the galactic plane. Among the galactic sources detected so far is possible to find (ordered by number of detections by *Fermi*): Pulsars, supernova remnants, globular clusters, gamma ray binaries, novae and the Galactic Center.

- **Pulsars:** They are fast rotating neutron stars, emitting electromagnetic radiation periodically. They are very compact ($\sim 12 - 13$ km in size) and dense (range between $\sim 1.4 - 3 M_{\odot}$) objects, resulting from the gravitational collapse of a massive star. As parent stars are rotating, momentum is conserved, scaling their rotation to huge velocities with spinning periods ranging between ms up to some seconds, producing an extremely intense magnetic field in their vicinity. A large collection of pulsars has been detected in the

HE range (~ 150 [24]) but only one, the Crab pulsar [25, 26, 27], has revealed an underlying component in the TeV range. Several models attempt to describe their γ -ray emission, by defining different emission regions (see Fig. 1.4): the pulsar magnetosphere (through synchrotron emission and curvature radiation), the un-shocked relativistic wind, and the synchrotron nebula (both through inverse Compton). The very high energy component of the Crab pulsar supported models which assumed the emission region is located in the outer magnetosphere. It must be noted High Energy Stereoscopic System (H.E.S.S.) recently detected the Vela pulsar, but no TeV component was observed [28].

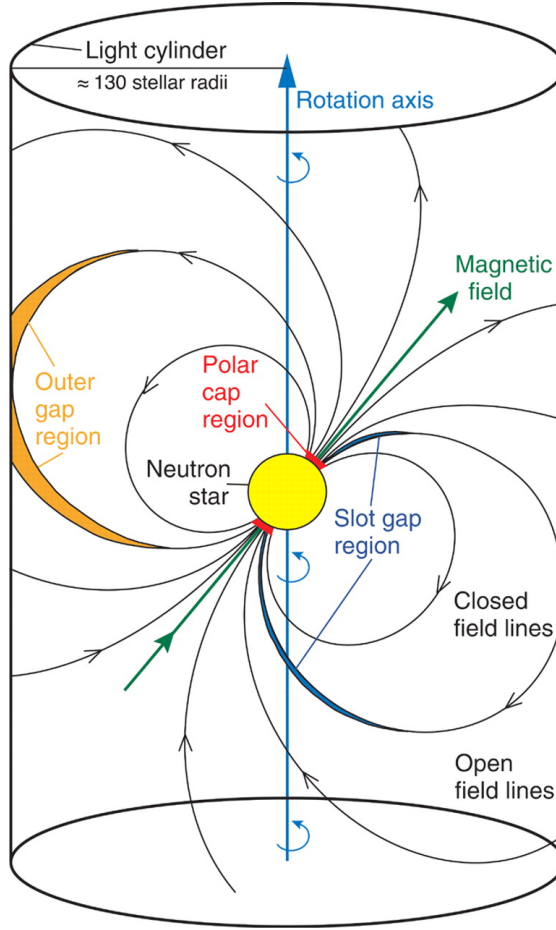


Figure 1.4: Illustration depicting the three proposed regions of γ -ray emission within the light cylinder of a pulsar: the Polar cap region (*ref*), the Slot gap region (*blue*) and the outer gap region (*orange*). Extracted from [25].

1. HIGH ENERGY ASTROPHYSICS

- **Supernova remnants:** A Supernova Remnant (SNR) is the structure resulting from the explosion of a dying star (supernova). The ejected material expands forming a shock wave against the interstellar medium. SNRs are divided into three classes, characterized by their observational features: *pulsar wind nebulae*, *shell-type* and *composite* remnants. *Pulsar Wind Nebulae (PWN)* are dominated by the presence of a central pulsar as a result of the supernova explosion. It emits a constant flux of relativistic particles referred as *wind*, mainly caused by its huge magnetic field. These particles interact with the medium forming a shock, losing energy through synchrotron, emitting electromagnetic radiation in the MeV range. They also produce inverse Compton scattering with low energy photons present in the shock, and with the MeV photons produced through synchrotron (self inverse Compton), producing a significant component in the GeV range. *shell-type* remnants have no central pulsar, or at least their winds are not strong enough to fuel the expanding shock, so the only observed characteristics are the ones produced by the hot shell, energizing particles through Fermi acceleration processes [29]. The third class, the *composite remnants*, is defined as a mixture of both described types: After the explosion of a massive star, both a shell-like expanding shock may be formed, with pulsar powered PWN located inside.

The Fermi acceleration processes that occur in the expanding shocks are believed to be the cause of most of the generated Galactic cosmic rays. These cosmic rays produce a large amount of hadronic interactions within the remnants, producing neutral pions, and therefore γ -rays. These have been observed by the *Fermi*-Large Area Telescope (LAT) [30], partially solving the mystery of the cosmic rays origin.

- **Globular clusters:** These structures are composed by a spherical dense distribution of stars, with a strong gravitational bond, orbiting a galactic core as a satellite. Located in the galactic halo, these systems were able to evolve earlier, which makes them very old structures. As stellar densities are high, collisions occur, which produced a large amount of pulsars and binary systems [31], both known γ -ray emitters. These objects produce ultra-relativistic electrons that could fuel inverse Compton scattering of different

photon populations, as the ones from the CMB, starlight or thermal photons generated in the center of the globular cluster.

These objects have been detected by *Fermi*-LAT (15 associated sources [24]) in the HE range. H.E.S.S. holds the achievement of detecting the first (and last, by the time this work was written) of the observed globular clusters in the VHE range, detecting Terzan 5 [32].

- **Gamma ray binaries:** Binary systems are formed by two gravitationally bound stars orbiting around a common center of mass. If one star is a compact object (either a neutron star or a black hole) and the other is massive enough to fill its Roche lobe, mass will be accreted by the compact one. The in-falling matter from the donor star releases gravitational potential energy as X-ray emission. Also γ -ray emission have been reported from these objects, with an still unclear origin. One possibility attributes the γ -ray emission to be similar to the one observed in quasars (see section 1.3.1.2), with the accreting matter and the central compact object creating jet-like emission [33, 34], naming them *micro-quasars*. The second scenario considers, the compact objects could be pulsars creating wind shocks similarly as the ones observed in SNRs [35].

Some of these binary systems have been detected in the HE range by *Fermi*-LAT [36, 37] and also by the current generation of ground detectors, such as H.E.S.S. [38, 39] and Major Atmospheric Imaging Cherenkov (MAGIC) [40, 41].

- **Novae:** As described for the gamma-ray binaries, matter is sometimes accreted between neighbouring stars. In systems with a compact accreting white dwarf (WD) with a close by companion, gasses build up in the compact star's surface, increasing their temperature. If the temperature of the accreted gas surpasses the proton-proton fusion temperature ($\sim 15 \times 10^6$ K), a sudden cataclysmic explosion called *Nova* may occur, increasing the star brightness for long periods of time (\sim months) and ejecting a large fraction of the gas accumulated in the WD's surface. These thermonuclear bursts of energy were never thought to emit up to the γ -ray range until *Fermi*-LAT

1. HIGH ENERGY ASTROPHYSICS

detected several of these events emitting at energies above 100 MeV [42]. Recent high resolution observations observed symmetries with the rotation direction: rapid gas ejection in the rotating poles while slow denser material was expelled along the equatorial plane. Synchrotron emission was found in their frontier, indicating a shock structure which could accelerate relativistic particles, capable of producing HE γ -rays [43].

- **Galactic Center:** The center of the Milky Way, the Galactic Center, is a very crowded region in the γ -ray band, believed to hold a super-massive black hole (Sgr A*), together with other γ -ray emitters such as SNRs as Sgr A East and stellar clusters with enhanced star formation. Sgr A* emission has been an issue of discussion: Milky Way could be an AGN, with γ -rays originated in the base of the jets, or the emission could be attributed to an accretion disk formed around the super-massive black hole. Other theories suggest it could be due to the presence of a high concentration of millisecond pulsars [44], or even Dark Matter annihilation [45]. Sgr A* has been detected both by *Fermi*-LAT [46] in the HE range and by ground based detectors in the VHE range: by H.E.S.S. [47], MAGIC [48] and Very Energetic Radiation Imaging Telescope Array System (VERITAS) [49].

1.3.1.2 Extragalactic γ -ray sources

Although most of the HE photons reaching the Earth come from our galaxy, the γ -ray sky is filled with point-like extragalactic sources. In fact, these are the most common sources in this energy range, dominating the Fermi-large-sized telescopes (LST) catalog (see Fig. 2.5). These sources are mainly *Active Galactic Nucleus (AGN)*, but also *Galaxy clusters*, *Starburst galaxies* and the intriguing transient phenomena, the *γ -ray bursts*.

- **Active galactic nuclei:** These are the brightest steady sources in the universe. This emission is believed to be the result of the accretion of vast amounts of gas by a Central Massive Black Hole (CMBH) located in the center of the host galaxy. An accretion disk is formed surrounding the BH due to its huge gravitational attraction, pulling the host galaxy's gas and

star-like objects. Gravitational acceleration and friction increases the energy of these particles, up to relativistic regimes. Although many types of AGNs have been defined, they are usually classified into 2 broad types: *radio-quiet* and *Radio-loud*, the latter with higher γ -ray emission. *Radio-loud* emission, originated from the CMBH, is beamed perpendicular to the accretion disk plane forming two jet structures accelerating particles to ultra-relativistic energies, producing electromagnetic radiation in the whole range of the electromagnetic spectrum. Further classifications are applied depending on their Spectral Energy Distribution (SED). The *unified scheme* (see Fig. 1.5) proposes these differences do not arise from intrinsic properties of these sources, and rather come from the angle between the direction of the jets and the line of sight [50]. Blazars are defined as those AGNs with their jets pointed towards our direction. Between blazars, two main categories arise: the FSRQ, with broad strong optical emission lines, and the BL Lac, with weaker emission lines characterized by their fast flux variability and optical polarization.

As shown in Fig. 2.5, blazars are the most frequent steady γ -ray sources. *Fermi*-LAT already detected more than 1100 identified sources associated with BL Lac or FSRQ active galaxies. They are also extensively observed in the VHE range by ground based telescopes, constraining the size of the radiating region to less than the 20% of the gravitational radius of its CMBH (after IC310 observations by MAGIC [51]). Also, the detection of VHE gamma-ray spectra of high redshift blazars performed by H.E.S.S. [52], MAGIC [53, 54] and VERITAS [55], [56] placed strong upper limits on the EBL density, indirectly measuring star formation rates along the history of the universe. For further knowledge on AGNs and detection techniques in the VHE see [9].

- **Starburst galaxies:** These galaxies are characterized by a very high star formation rate, generally caused by the gravitational interaction between closeby galaxies (passing near or mergers). This starburst is just considered a phase of galactic evolution, as the gas producing star formation will be likely consumed in very short time scales (compared with galactic lifetimes). As a result of induced star formation, large quantities of massive stars are created

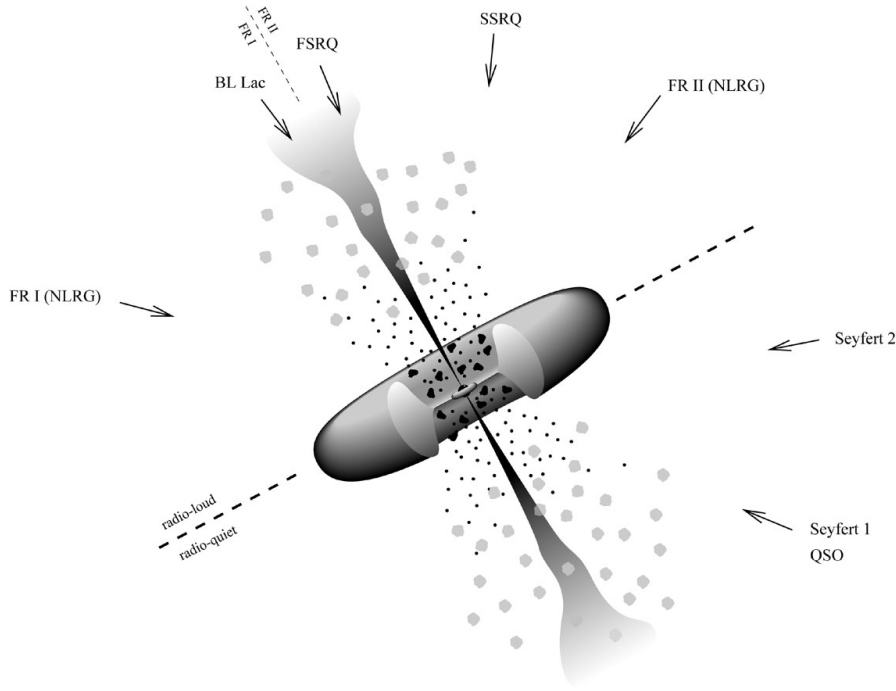


Figure 1.5: Sketch of the unified scheme of the AGNs. The vast majority of the γ -ray emitting AGNs are the blazars, those which their jets are pointed close to the line of sight. Extracted from [57].

producing higher frequencies of supernovae, with their corresponding SNRs. γ -ray emission is expected from the cosmic ray acceleration interacting with the medium through hadronic collisions, and from the SNRs shock fronts, producing inverse Compton scattering.

Several starburst galaxies have been detected in the VHE range and confirmed by *Fermi* (NGC 253 and M82 [58, 59, 60]).

- **Gamma-ray bursts:** Gamma-Ray Bursts (GRBs) are fleeting events considered the brightest sources of electromagnetic radiation in the observed universe. Extremely powerful bursts of γ -ray radiation are emitted in very short time scales, from few seconds for *short Gamma-Ray Burst (GRB)s* to few minutes for *long GRBs* or rarely hours for *ultra-long GRBs*. They are capable of emitting, considering isotropic emission, in the order of $10^{52} - 10^{54}$ ergs, equivalent to the whole radiation emitted by a galaxy like the Milky Way over a period of years. Their initial γ -ray burst (\sim MeV-GeV) is generally

followed by an *afterglow*, emitted at temporally decreasing energies: X-ray, ultraviolet, optical, infrared, microwave and radio. After their detection in 1967 [61], multi-wavelength observations with a long list of contributors over the whole electromagnetic spectrum allowed to measure their distances and provide valuable information of their nature: these events occur in cosmological distances, ranging between $z = 0.0085$ up to $z = 6.7$, with an average of $z = 2.3$ – 2.7 and are probably generated by ultra-relativistic outflows, being likely collimated.

The general picture is described as follows: The *Short GRBs* generally lasting less than 2 seconds, have been associated with evolved regions, with no star formation. The leading interpretation assumes these events are caused by the merging of very compact objects, such as neutron stars or black holes. The most common ones, *long GRBs* lasting more than 2 seconds, tend to show bright afterglows allowing a deeper study along different wavelengths. The vast majority of these events have been associated with regions of rich star formation, even type II supernovae [62]. These associations proof *long GRBs* are a different population of events, and are likely caused by massive stars explosions.

GRBs have been observed in the HE range by satellites, first by the Energetic Gamma Ray Experiment Telescope (EGRET) and later on by Fermi's Gamma-ray Burst Monitor (GBM) [63]. Attempts to observe this events in the VHE range have been performed by MAGIC, specifically designed to detect such objects, but only upper limits were possible for the time being [64, 65].

- **Galaxy clusters:** Galaxy clusters are the largest gravitationally bound structures in the universe (together with the superclusters) consisting of hundreds or even thousands of galaxies. The standard models of structure formation predict larger structures take longer to collapse, implying that galaxy clusters are still being formed. In this process, violent energetic events are bound to occur, expected to produce shock fronts capable of accelerating cosmic rays. These objects have been proposed as possible γ -ray emitters for several reasons: These shock fronts are expected to accelerate electrons

to relativistic energies, so inverse Compton scattering is expected to occur [66]. In addition, the cosmic rays created in these shocks and the added contribution of the contained galaxies (AGNs or SNRs) are expected to be confined within the cluster due to its magnetic field and bast scale structure, eventually producing hadronic interactions, and eventually γ -ray emission. No galaxy cluster has ever been observed in γ -rays, and only upper limits have been calculated [67, 68].

1.3.2 Potential WIMP annihilating targets

As introduced in section 1.2.2, diffuse γ -ray emission could be emitted by DM particles under certain considerations. As described by Eq. 1.8, the astrophysical factor J is the main parameter to take into account for DM searches. The best targets for such research are the following: galaxy clusters, the Galactic Center, or dwarf Spheroidal (dSph)s galaxies.

- **Galaxy clusters:** As previously described, these are the largest gravitationally bound structures in the universe. They have masses ranging between $10^{14} - 10^{15} M_{\odot}$, where DM is believed to account for the 80% of their mass budget. This is the reason why such sources where always considered as excellent targets for DM detection. In addition, N -body simulations seem to show DM distribution in these objects generally produces inhomogeneities, capable of boosting up the intrinsic WIMP annihilation flux by a factor 100 – 1000. For these reasons, the current generation of γ -ray detectors, both by satellites in the HE range and by ground based detectors in the VHE range have placed strong upper limits on self annihilating DM particles using these sources [69, 70, 71, 72]. It must be noted, as previously introduced, that galaxy clusters may have alternative processes capable of emitting diffuse γ -ray emission, so attributing a detection to DM annihilating particles could still be argued.
- **Galactic Center halo:** Recent simulations seem to show the Galactic Center (GC) halo would be the dominant source of DM annihilation within our reach. Unfortunately, as previously introduced, the GC is a crowded region,

where γ -ray emission have been already reported from several sources: Sgr A*, the PWN G 359.95–0.04 [73] and galactic diffuse emission, consistent with hadronic collisions of cosmic rays in molecular clouds. For this reason, DM searches should focus on "cleaned" regions, as close as possible from the GC, but without the contribution of these γ -ray known sources. Recent observations by H.E.S.S. using this approach, comparing 112 hours of data taken at 0.3° and 1.0° from the GC have placed the strongest upper limits on WIMP annihilation above 100 GeV. *Fermi*-LAT yields the best upper limits in the HE range up to date [74], stacking the expected signal from the GC halo and from the dwarf galaxies.

- **Dwarf spheroidal galaxies:** These dwarf galaxies are small and faint galaxies located in the Milky Way halo. These sources have very high mass to light ratios, and stellar dynamics show they may be among the most DM dominated known sources. In addition, N -body simulations predicted the presence of DM clumps populating the Milky Way halo [18]. Under the Λ CDM assumptions, these dSphs are proof of the presence of the DM clumps. With approximately known astrophysical factors (J) and practically background free, these sources are ideal targets for DM annihilation prospects, as a significant excess in γ -rays consistent with an expected annihilation shape would mean the first unequivocal DM detection. These sources have been profoundly studied by the current generation of γ -ray telescopes, both from ground and space [75, 76, 77, 78, 79, 80]. It must be noted that the expected "boost factor" in dSphs is significantly smaller than in other cases, being $\mathcal{O}(1)$ [81].

Chapter 2

Past, present & future of γ -ray detectors

As far as we can go into the past, humankind used astronomy. First, direct naked eye observations were performed to understand the effects of seasons, the length of a day, years periodicity, or to inspire ceremonial rituals. Ancient civilizations had already some kind of dedicated astronomical observatories, such as the ones found in the Middle East, China, Greece, India or Central America. Throughout our history fascination over the unknown drove thinkers, philosophers and intellectuals into the search of an answer, an explanation that could bring logic to their surroundings.

The first mayor revolution astronomy underwent was the invention of the telescope, in the beginning of the *XVII* century. A cascade of discoveries followed, directly leading to Newton's law of universal gravitation. From that moment on, all major discoveries in astronomy were direct consequences of the development of new techniques, tools or methods that improved sky observation instruments. In the early *XIX* century, the invention of photography and spectroscopy opened a vast window of knowledge. Astronomers could now study star's composition, mass, temperature or size.

But it was in the *XX* century that modern astronomy opened a whole new universe to scientists all around the world. With the detection and study of exotic objects such as quasars, pulsars, blazars, radio galaxies, neutron stars or black

holes, a stream of theories emerged to explain their existence and observed properties. Specially after the end of the second world war, a long list of different experiments broadened the observational window, expanding the narrow band of the optical wavelength down to the infrared, microwave and radio, and up to ultraviolet, X-ray and γ -ray.

This chapter aims to explain the different detection techniques used in the High Energy astronomy and the latest major experiments. In section 2.1 historical motivations leading to past and current experiments in γ -ray astronomy will be discussed. In section 2.2 the different techniques developed to observe gamma-rays are presented, explaining the different complications each energy range needs to overcome. Section 2.3 shows an overview of current gamma-ray experiments and finally, in 2.4, the future γ -ray observatories to come are introduced.

2.1 Historical motivations

Historical motivations leading to the first γ -ray detectors are closely related to cosmic ray astronomy. Although the distinction between cosmic ray and gamma ray was not clear at the time, the similar nature and detection mechanisms used for high energy nuclei and high energy photons made this two scientific topics grow intertwined.

After the invention of the first electrometer by Theodor Wulf [82] in the beginning of XX century, a set of different measurements of atmospheric ionization were undertaken. This ionization was first associated with Earth's radioactivity, but studies by Domenico Pacini [83] and later on by Victor Hess [5] showed an increasing ionization level with altitude. This experimental evidence demonstrated that such radiation had outer space origin. Victor Hess received the Nobel Prize in Physics in 1936 for his discovery.

The term "cosmic rays" was defined after Robert Millikan [4] improved V. Hess measurements extending the study to high altitudes and under water. Initially, the origin of this radiation was assumed to be electromagnetic, until a dependence with latitude was measured, showing interaction with Earth's geomagnetic field. It was then concluded that cosmic rays were charged, and could not be photons. Subsequent discoveries lead to a better understanding of cosmic rays nature, but

their origin was still unknown.

During the mid-XX century, some theoreticians (Henry Primakoff [84], Sachio Hayakawa [85] or P. Morrison [86]) predicted the existence of certain processes in the universe, such as supernova explosions or cosmic ray interactions with interstellar gas, that could produce high energy photons. Paradoxically atmospheric absorption, the very same phenomenon that lead to cosmic rays discovery, prevented γ -ray detection inside Earth's atmosphere.

Direct γ -ray detection took advantage from the spacecraft technology boost after the Second World War and deployed the first γ -ray telescope into orbit, carried by the Explorer 11 satellite in 1961. Subsequent missions brought great success in the field, specially SAS-2 (1972 [87]) and COS-B (1975-1982 [88]) satellites, creating a γ -ray source catalogue with 25 sources [89] and a map of the Milky Way in this energy range, shown in Fig. 2.1.

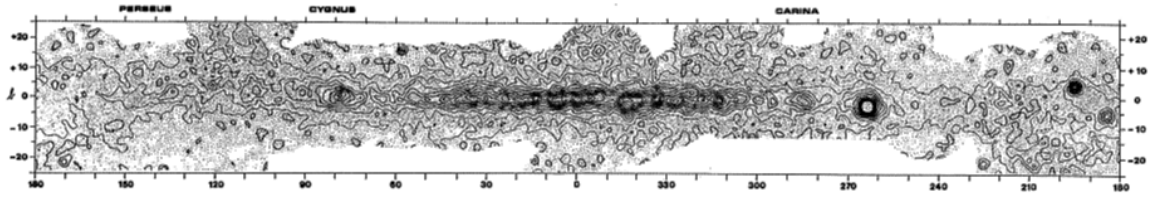


Figure 2.1: Cos-B Milky Way map in the γ -ray range.

In the late 1960s a different pioneering technique was developed by the Whipple collaboration for the observation of VHE γ -rays from ground: Imaging Atmospheric Cherenkov Telescope (IACT), resulting in the detection of the first TeV γ -ray source, the Crab Nebula, in 1989.

2.2 Observation Techniques

As shown in figure 2.2, the atmosphere creates a natural barrier protecting Earth's surface from photons with energies higher than the visible light range ($\lambda < 390 \text{ nm}$), preventing direct γ -ray detection from the ground. Consequently, direct γ -ray detection must take place above the atmosphere, either by satellites or balloons. These detectors generally observe the low and medium energy range (up to around

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

100 GeV). They can not access higher energies due to their limited effective area and the low fluxes observed at these energies.

A different technique was developed to observe higher energy ranges. It consists in the observation from ground of the secondary products: particles or radiation produced in the cascades generated by the collision of a high energy photon or atomic nuclei. This technique permits ground based detectors with huge collection areas, allowing the observation of the very, ultra and extremely high energy ranges (with energies above ≈ 20 GeV).

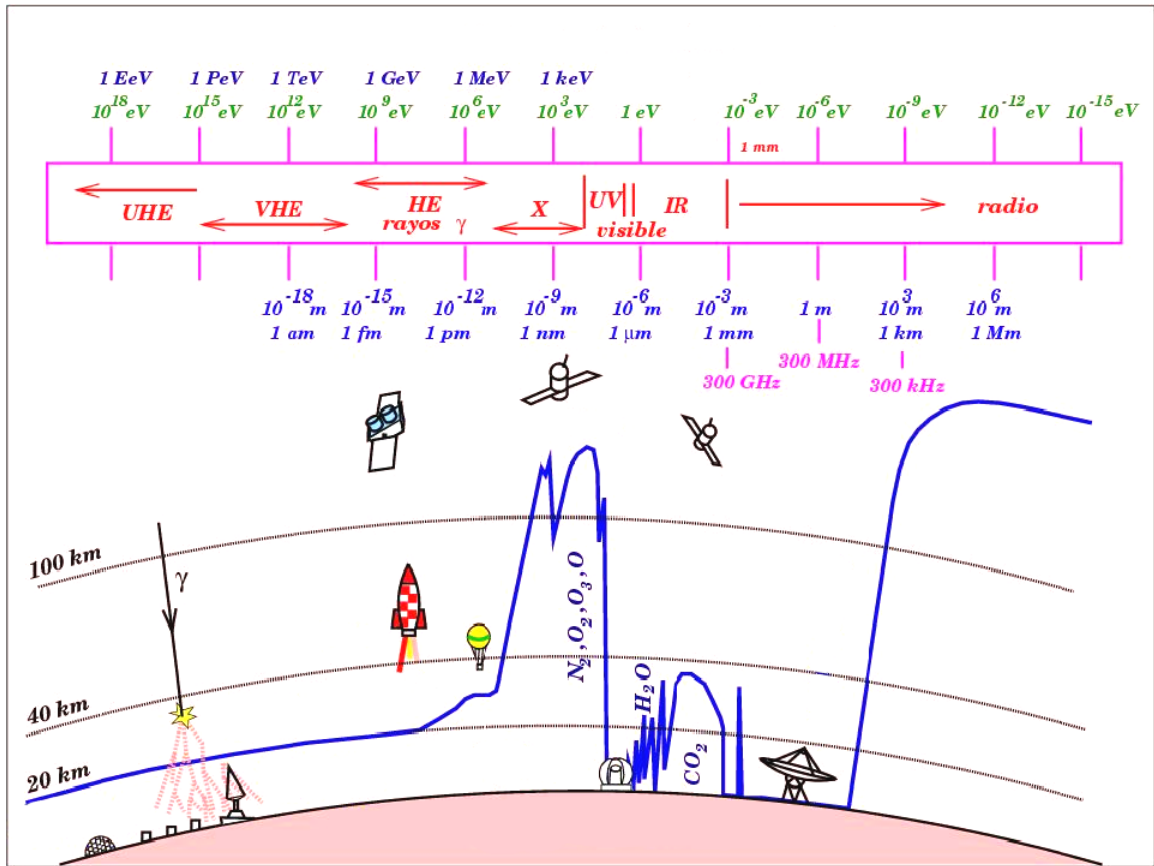


Figure 2.2: Altitude where 50% of the electromagnetic radiation is absorbed by the atmosphere (blue line) along the different energy bands of the spectrum. Different techniques used are pictured, including ground, air and space observations.

2.2.1 Space detectors

High altitude balloons and space satellites are the only experiments able to directly observe high energy cosmic photons. These instruments carry particle physics detectors, placed beyond the atmospheric absorption. Depending on the energy range of study, the observation technique is based on different photonic interactions.

In the low end of the γ -ray band, the detection is based on the γ -ray Compton scattering of electrons or the photoelectric effect. The typical Compton detector consists of two planes: after the γ -ray collides with the first plane, measuring its crossing position, a deflection of its trajectory is produced by Compton scattering with atomic electrons in the medium. On arrival to the second plane, the final position is recorded and the γ -ray is absorbed by a calorimeter where its energy is measured. Compton effect formulae allow to reconstruct to some extent the arrival direction and the energy of the primary γ -ray out of the deflected direction and calorimeter measurements. The IBIS spectrometer aboard the INTEGRAL satellite [90] exploits this technique.

In the HE range, the main detection principle is the pair production. High density materials are used to force γ -rays to annihilate into electron-positron pairs, which are then measured to estimate impinging γ -ray properties. Detectors usually consist on a tracker built over a calorimeter, surrounded by a plastic anti-coincidence detector. The tracker is composed of layers of a high density material interleaved with particle detectors. Incoming γ -rays go through the plastic anti-coincidence detector freely, while cosmic rays cause a flash of light, identifying these as background. The γ -rays continue until they interact with an atom in one of the high density layers (generally tungsten foils) producing electron positron pairs. These particles interact with the tracking detectors creating ions on their way until they are stopped by the calorimeter, measuring the total deposited energy. Combining the measurements of the anti-coincidence detector, the tracker and the calorimeter, the energy and the original direction of the γ -ray are reconstructed, while background is suppressed. The LAT instrument on board of the Fermi Gamma-ray Space Telescope is the latest example of this observation technique, and will be discussed further in section 2.3.1.

2.2.2 Ground detectors

Current VHE and Extremely High Energy (EHE) γ -ray astronomy is driven by ground based detectors. Instead of interacting with incident γ -rays like space detectors, these instruments measure the products generated by the collision of very high energy γ -rays striking the atmosphere. These interactions produce cascades of particles. Among them, those that are charged and travel with a velocity larger than the local speed of light emit Cherenkov radiation. These showers of particles are called EAS and are explained with a deeper scope in Appendix A. The measurement of the EAS properties, with different techniques depending on the energy range, allow the reconstruction of the direction and the energy of the original particle. The major problem affecting this technique is the huge background of cosmic rays, as they generate similar showers as γ -rays. Understanding intrinsic differences between EAS from different primary particles becomes crucial for background suppression.

Depending on the energy range of interest, current γ -ray detectors use different techniques:

- **Atmospheric Cherenkov detectors:** These instruments measure the Cherenkov light produced in EASs. There are two different detection approaches currently in use: *sampling detectors* and *imaging detectors*.

Sampling detectors consist on a grid of counters measuring arrival time and density of Cherenkov light arriving from the front of EASs over a wide region on the ground. Using the intensity and relative times recorded on the different detectors, the primary particle original direction and energy can be reconstructed. The main disadvantage of these detectors resides on the poor information acquired from the development of the EAS, therefore differentiating γ -rays from background becomes problematic. Nowadays imaging detectors outperform this technique, and their contribution to the field is not very significant.

Imaging detectors, instead of sampling the front of the EASs, create an image of the shower as classical optical telescopes. With these snapshots is possible to study the evolution of the EAS along the atmosphere, and parametrize dif-

ferent properties of the cascade. These parameters allow to differentiate between γ -ray and cosmic ray induced EASs, greatly improving signal to noise ratio, and therefore sensitivity. The first Imaging Atmospheric Cherenkov Telescope (IACT) able to detect an astronomical source was Whipple [91], in operation since 1968, detected the Crab Nebula in 1988. Nowadays, VHE Astronomy is driven by the current generation of IACT such as the MAGIC [92], the H.E.S.S. [93] and the VERITAS [94]. See section 3.1 for more details about IACTs.

- **Water Cherenkov detectors:** This technique also measures Cherenkov photon production, although in this case in a different medium with higher refractive index, increasing photon emission. Grids of opaque water tanks are spread over a wide area, measuring the Cherenkov emission from particles generated in EASs. As particles need to enter the tank before they are absorbed, these detectors need to be located at very high altitudes. This technique uses no optical focusing so the Field of View (FoV) of the detectors is significantly larger than IACTs, although at the same time they have a worse angular resolution, as direction reconstruction uses only timing and the intensity of the footprint. Detectors can be built as a single water pool, like MILAGRO experiment [95], or a grid of smaller individual water tanks, improving resolution, like the future HAWC [96] Observatory. HAWC which be analyzed with a deeper scope in section 2.4.1.
- **Particle counter matrices:** These detectors directly measure EASs induced particles using matrices of counters. These counters use classic particle detectors recording the arrival time and direction of the incoming particles, and deriving from them the energy and direction of the primary. As an example the Tibet-AS [97] experiment covers an area of 36.000 m² with 697 individual scintillation counters. The Argo-YBJ [98] experiment uses a full coverage detector consisting of a single layer of resistive plate counters. These detectors are able to observe in the VHE range, although their sensitivities are not as competitive as current IACTs or water Cherenkov experiments.
- **Fluorescence detectors:** Cherenkov emission is not the only source of

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

photons in EASs. There is a fluorescence isotropic emission along EAS proportional to the energy of the impinging particle. In the atmosphere, charged particles generated by EASs excite nitrogen molecules, which eventually decay producing photons in the ultra-violet and visible range. Detectors are able to collect this component from very long distances collecting the fluorescence light with mirror systems directed towards Photo-Multiplier Tube (PMT) detectors. When several detectors observe the same shower in coincidence, they can measure the intensity and time of arrival of the radiation, reconstructing the development of the shower along the atmosphere, inferring direction and energy of the primary cosmic ray. Taking into account the low efficiency of fluorescence emission in the atmosphere, only cosmic rays in the Ultra High Energy (UHE) and EHE range are observed with this technique. This technique has been successfully applied in the High Resolution Fly's Eye (HiRes) experiment [99], studying the composition of UHE cosmic rays.

- **Hybrid detectors:** These detectors take benefit from several of the previously listed techniques at the same time. Each technique can be used to study different aspects of an EAS which then are combined. The Pierre Auger Observatory [100], combines 1600 water Cherenkov detectors with 27 optical telescopes to measure fluorescence from UHE cosmic rays, observing energies beyond $10^{18}eV$ over an area of $3.000Km^2$.

2.2.3 Ground vs Space

All the detection techniques discussed over sections 2.2.1 and 2.2.2 have certain advantages and disadvantages, specially affecting the energy range they are able to observe.

Direct detection of gamma and cosmic rays by space detectors allow a wider field of view with excellent background rejection. Their main limitation is the high costs associated with sending payloads into orbit, constraining the size and weight of the instruments on board, restricting collection areas to few m^2 . Besides, the accuracy of the directional reconstruction of current trackers is limited by the multiple scattering caused by the thickness of the foils. Thicker layers increase

effective area for high energy γ -rays although at the same time increment multiple scattering effect, broadening the Point Spread Function (PSF).

On the other hand, construction of ground detectors permits lower budget projects, without important size/weight restrictions. Additionally, EASs extended nature increases collection area in several orders of magnitude compared with space detectors, extending the observable spectrum to higher energies. The use of stereoscopic observation allows accurate direction and energy reconstruction and a good background rejection. The main drawback of these techniques is the overwhelming background rate caused by cosmic-rays and their similar Cherenkov emission than γ -rays generated EAS within the HE range.

These limitations divide the whole gamma ray band into different energy ranges, shown in table 2.1:

- **Below ≈ 50 GeV:** Corresponds to the low end of the γ -ray band. Needs to be observed from space, as EAS generated by incident photons in the atmosphere are not energetic enough to produce sufficient Cherenkov radiation. The Low Energy (LE) band corresponds to energies below 30 MeV, only detectable by the photoelectric and Compton effects, as they are not energetic enough to produce electron-positrons pairs. The energy range between 30 MeV and 50 GeV photons is the HE range and is currently driven by pair production detectors.
- **50 GeV to 30 TeV:** Is labeled as the VHE range. IACTs are currently the most sensitive instruments in this range as a result of their augmented collection areas and decent background rejection. Above ≈ 300 GeV EASs are energetic enough for particles to be detected by water Cherenkov detectors placed at very high altitudes.
- **30 TeV to 30 PeV:** UHE photons generate EAS with their shower maximum at lower altitudes, where Water Cherenkov Detectors can be placed. As water tanks are opaque, they are fully operational during daytime greatly improving observation times compared with imaging detectors, measuring a big fraction of the observable sky at the same time. Their wide field of view and high duty cycle turn these detectors to be the most efficient to detect the low frequent UHE photons.

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

- **Above 30 PeV:** The EHE range corresponds to the most energetic side of the γ -ray band. The main limitation on sensitivity is the extremely low frequency of events, therefore observatories need to cover very wide areas. Fluorescence radiation becomes significant at these energies, it's isotropic emission allows individual events to be observed from very long distances ($> 10^4$ m), turning it into the most viable detection technique at these extreme energies.

Table 2.1: Different γ -ray observational subranges.

Range	Notation	Detection mechanism	Type of detector
< 30 MeV	LE	Photoelectric, Compton	Space-based
30 MeV-50 GeV	HE	Pair-production	Space-based
50 GeV-30 TeV	VHE	Cherenkov (atmosphere)	Ground-based
30 TeV-30 PeV	UHE	Cherenkov (water)	Ground-based
> 30 PeV	EHE	Fluorescence, hybrid	Ground-based

Convenient segmentation of the the γ -ray band, together with the corresponding type of detector and detection mechanism currently used for the exploration of each subrange. LE, HE, VHE, UHE, and EHE stand for low, high, very high, ultra high, and extremely high energy. Adapted from [10].

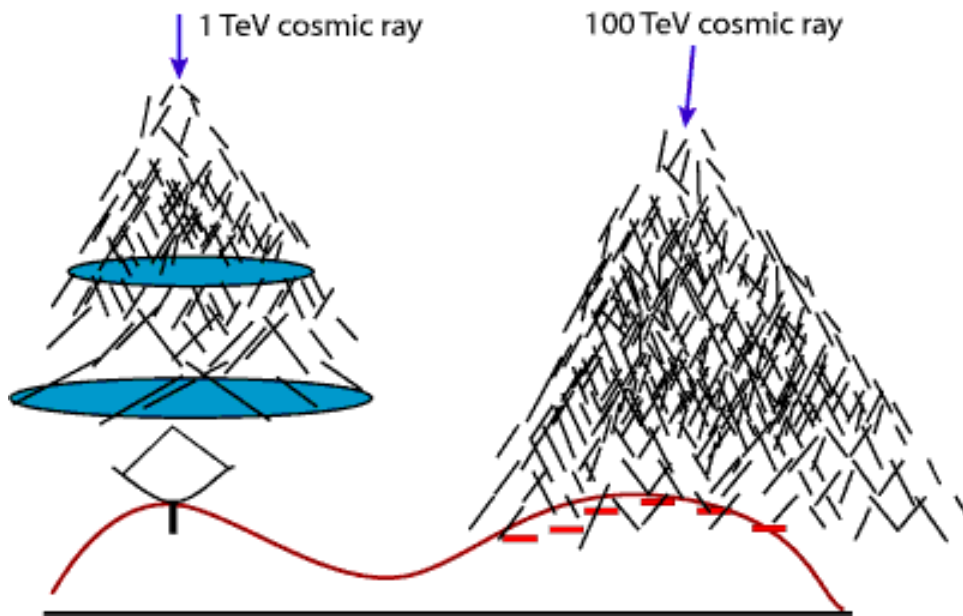


Figure 2.3: Extended Atmospheric Shower detection techniques at two different energy ranges. Generated Cherenkov light is observed by IACTs (left) in the VHE range while produced particles in the UHE range are observed by water Cherenkov detectors (right). Image courtesy of Milagro collaboration (University of California, 2002).

2.3 Current γ -ray observatories

Currently, the most sensitive experiments, leading the VHE astronomy field are **Fermi Gamma-ray Space Telescope** and the current generation of IACTs: **MAGIC**, **HESS**, and **VERITAS**. Their sensitivity is shown in Fig. 2.4 together with the one of **HAWC** and **MILAGRO**, not operative since April of 2008.

2.3.1 Fermi Gamma-ray Space Telescope

The Fermi Gamma-ray Space Telescope [103] is a space observatory designed to perform an all sky survey of the observable universe in the HE range. Launched in 2008 and placed in a low Earth orbit, it sweeps the whole sky every 3 hours, corresponding to two complete orbits around Earth. The 2 principal instruments on board of the satellite are the *Large Area Telescope (LAT)*, it's main instrument intended to perform the survey, and the *Gamma-ray Burst Monitor (GBM)*,

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

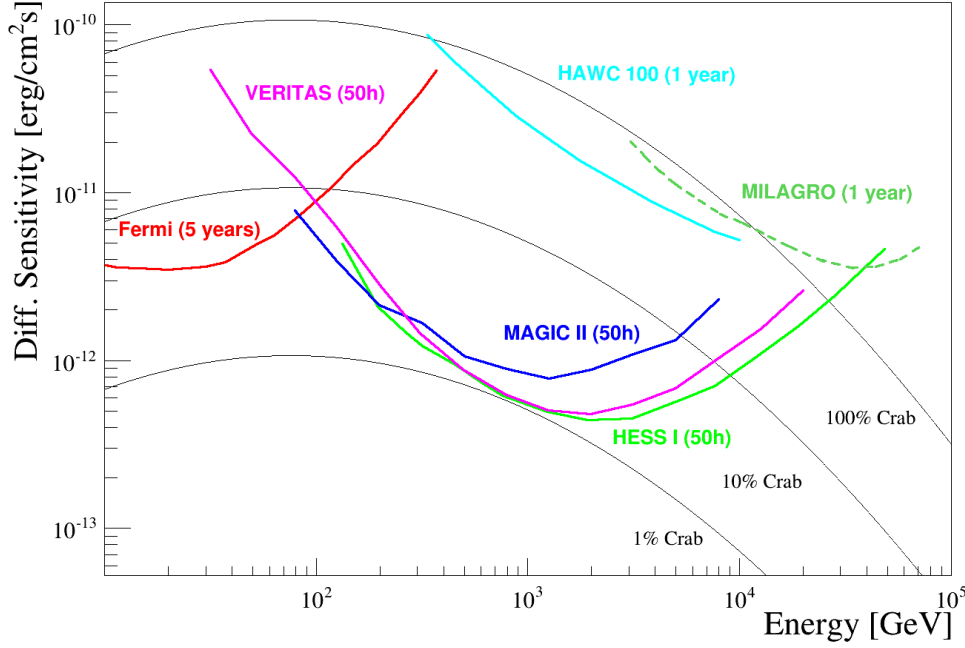
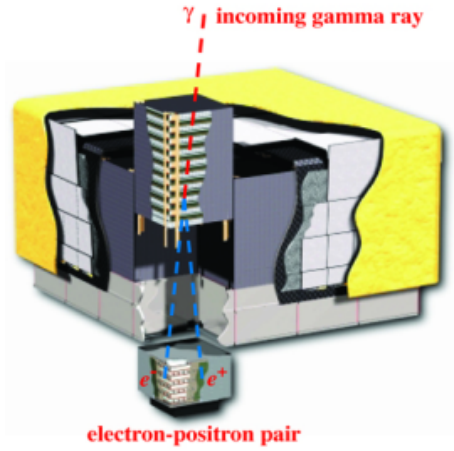


Figure 2.4: Minimum detectable flux of a point-like source for present VHE gamma-ray detectors. IACTs sensitivities correspond to 50h of observation while Fermi, HAWC (still in construction phase) and Milagro (currently not operative) correspond to sky surveys through several years. Dotted lines correspond to the 1, 10 and 100 % of the Crab Nebula, the standard candle in γ -ray astronomy. Data obtained from [101], [102], [94] and private communication with K. Bernlöhr.

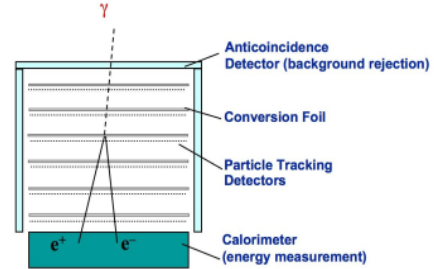
designed to monitor and study GRBs. Although the initial mission lifetime was planned to be 5 years, on August 2013 it was extended for another 5 years.

Within the key scientific objectives of the mission were the study of the γ -ray sky and sources, emission processes of AGN, pulsar and SNR, study and monitoring of GRBs, and the search for Dark Matter signals in the Galactic Center.

Data collected by Fermi mission is public, open for the community, and the collaboration publishes periodical catalogs gathering all statistically significant high energy sources detected so far.



(a) Fermi-LAT cut view.



(b) Individual "tower" schematic view.

2.3.1.1 Instruments on board

- **Large Area Telescope:** The LAT [104] is a gamma-ray detector able to observe incident γ -rays between 30 MeV and 300 GeV. It consists of a 4×4 grid of "tower" pair-conversion detectors (see a section sketch of the detector in figure 2.5a) wrapped with a plastic anti-coincidence scintillator responsible for background suppression.

Each "tower" is a converter-tracker detector consisting of 18 tungsten layers intertwined with 16 silicon trackers, built over a calorimeter module. Electron-positron pair production takes place in the tungsten layers while silicon-strip detectors measure the position of the e^\pm pair (see schematic view in figure 2.5b). The direction of the incident particle is reconstructed using these tracks. The angular resolution of the LAT (shown in figure 2.8) varies by an order of magnitude from the low to the high energy range. It is aided by the high precision of the silicon trackers resolution and limited by multiple scattering effect. The calorimeter modules contain a total of 96 narrow crystal scintillators distributed in 8 layers. Incoming electron-positron pairs produce electromagnetic showers generating scintillation photons, which are measured. Crystals are alternating in orientation so deposited energy can

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

also be tracked, increasing both energy estimation and background rejection.

The anti-coincidence detector is made of 89 individual plastic sections which allow γ -rays to pass undetected, while charged particles (such as cosmic-rays) produce flashes of light. Events are suppressed if the individual plastic scintillator corresponding to the incident particle direction was triggered.

The data acquisition system implements a multi-level trigger and provides an on-board event processing system running filter algorithms to reduce the data before space-Earth transmission. It also provides a science analysis platform to rapidly search for transients.

- **Gamma-ray Burst Monitor:** The GBM is a detector constantly searching for bright transient events, mainly GRBs. It includes 2 sets of scintillators, Sodium Iodide (NaI) and Bismuth Germanate (BGO) which cover respectively from a few keV to about 1 MeV and from ~ 150 keV to ~ 30 MeV. The NaI scintillators provide the burst trigger and location while BGO detectors grant an overlap between NaI and LAT energy ranges.

Data is buffered with a $5\mu s$ resolution, recording $\sim 50s$ of pre-trigger measurements. Any significant increase in the count rate from the NaI scintillators trigger data acquisition, and GBM calculates the preliminary position and spectrum of the source, and sends an autonomous re-point recommendation to the satellite.

2.3.1.2 Main results

The Fermi mission has produced many important results revolutionizing our knowledge of high-energy gamma ray sources and providing a huge amount of legacy data open to the community. With over four years of collected data, the Third *Fermi* LAT Catalogue (3FGL) [24] lists a total of 3033 point-like sources showing a crowded map with a wide variety of different source types, as shown in Fig. 2.5, including 1009 still unassociated and therefore still with unclear origin. In section 5.3 this problem is faced and Machine Learning Algorithms (MLA) are proposed to assess source type determination using the Second *Fermi* LAT Catalogue (2FGL) data. As proof of the success obtained by Fermi, the 2FGL publication is the most

cited article in astrophysics of 2012, showing the great impact and current bloom γ -ray astronomy is undergoing in the last decade.

Within the main scientific discoveries of the Fermi mission there are key milestones like proving the origin of galactic cosmic rays [30], new source discoveries such as pulsars with no counterparts in other wavelengths [105], the most energetic GRB ever observed [106], and completely unexpected detections as the so called "Fermi Bubbles", two massive diffuse gamma-ray and x-ray structures found around the center of the Milky Way, and never observed before [107]. In addition Fermi-GBM detected a total of 1310 GRBs, and 73 of them were also seen by LAT, extending their observed spectrum up to ~ 300 GeV.

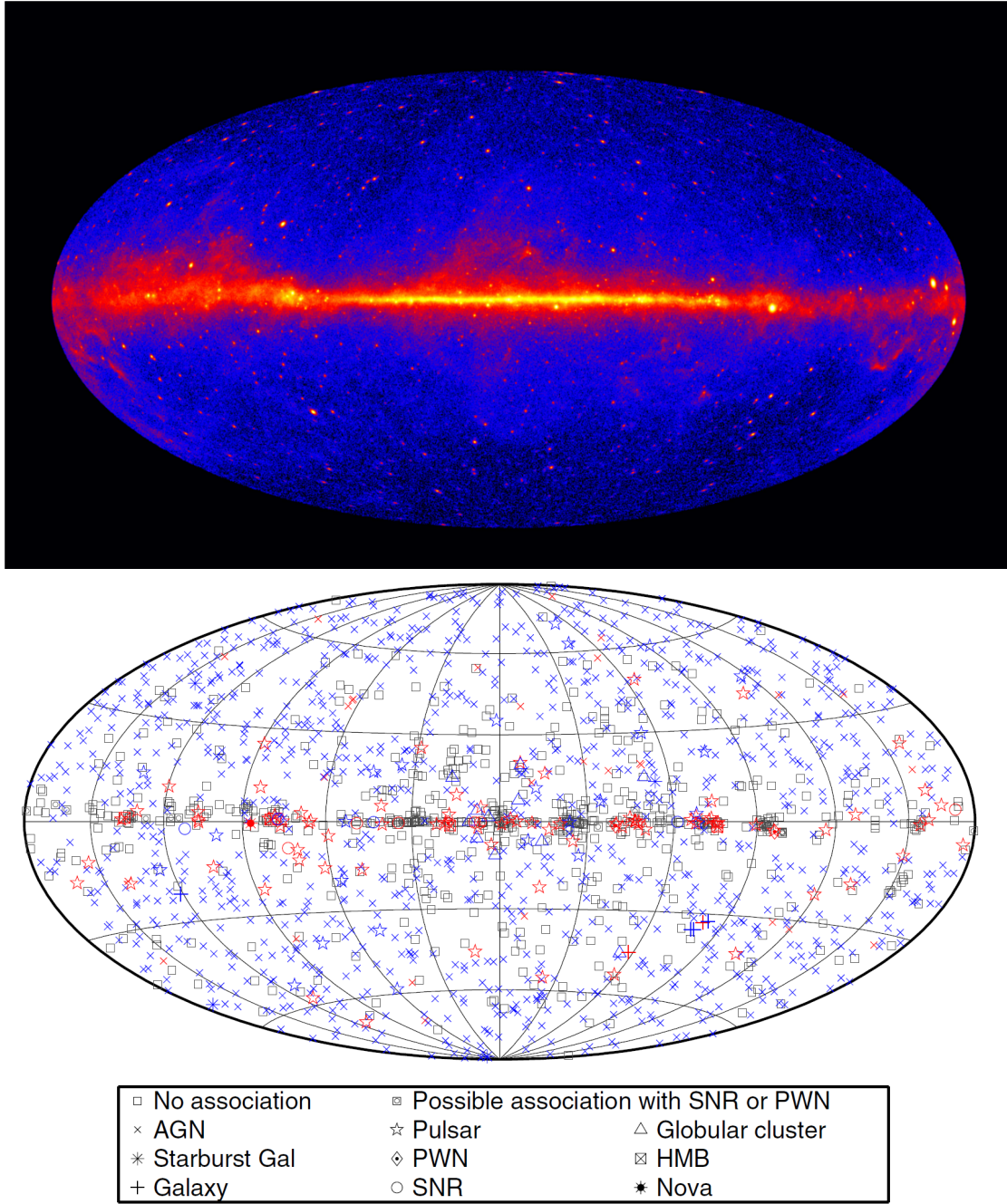


Figure 2.5: Full sky map in galactic coordinates showing gamma rays with energies over 1 GeV observed by Fermi-LAT along 5 years of operation (top) and point like sources present in the 2FGL, illustrating the different source types (bottom). The top image is courtesy of the Fermi collaboration [108] and the bottom image has been extracted from [109].

2.3.2 Imaging Atmospheric Cherenkov Telescopes

After the second half of the *XX* century, many experiments attempted to apply the IACT technique in astronomy. The first detection was carried out by the Whipple telescope, detecting the Crab Nebula [110] after 20 years of operation. A few years later stereoscopic imaging was developed by the HEGRA [111] collaboration, greatly improving IACTs sensitivity. All competitive experiments from the current generation of IACTs, shown with their specifications in Table 2.2, incorporate this technique with different approaches, such as increasing the individual telescope size or the total number of telescopes used. Nowadays IACTs provide the best sensitivity among the techniques used in the VHE range. The IACT technique and analysis will be explained in more detail in Sec. 3.1.

Table 2.2: Specifications of the different generations of IACTs

Instrument	Lat.	Long.	Alt.	Telescopes		Construction	FoV	E_{th}	ΔE	$\Delta\Omega$	Sensitivity in 50h
				Units	Area						
	[°]	[°]	[m]		[m ²]	[year]	[°]	[GeV]	[%]	[°]	[% Crab]
Whipple	32	-111	2300	1	75	1968	2.3	300	30	0.1	15
HEGRA	29	18	2200	5	8.5	1987	4.3	500	15	0.1	5
CAT	42	2	1650	1	17.8	1996	4.8	250	20	0.14	15
H.E.S.S.	-23	16	1800	4	107	2003	5	100	15	0.1	0.7
H.E.S.S. II	-23	16	1800	1	616	2012	3.7	20	10-40	0.3	0.7
MAGIC	29	18	2225	2	234	2004	3.5	50	15	0.07	0.8
VERITAS	32	-111	1275	4	106	2007	3.5	100	15	0.1	0.7
CTA (North)	30	–	~ 2000	~ 30	400/100	2017-2020	5-8	30	7	0.03	0.05
CTA (South)	-25	–	~ 2000	~ 125	400/100/15	2017-2020	5-8	30	6	0.03	0.05

Specifications of past, current and future generation of IACTs. Showing site location (latitude, longitude and altitude), total number of telescopes and individual mirror surface, date of first light, FoV, low energy threshold, energy and angular resolution and sensitivity (in the most sensitive energy, in Crab Units). Adapted from [112].

All IACT experiments share certain key scientific objectives. Some examples are the observation of SNRs, considered to be the main accelerators of cosmic rays, understanding the emission processes in γ -ray pulsars and discerning between proposed models, observing AGNs to understand the physical processes taking place in the vicinity of their super massive black holes and measuring the absorption affecting γ -rays from distant extragalactic sources due to the EBL, or searching for Dark Matter signals from regions where it is expected to accumulate such as

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

the Galactic Center or Dwarf Spheroidal Galaxies and detecting the high energy end of GRBs.

IACTs observable energy range overlaps with γ -ray space telescopes. First, with EGRET results, and later on in coalition with Fermi-LAT, IACTs provide improved angular resolution to classify Fermi unidentified sources (generally caused by the superposition within Fermi's wide PSF of several possible sources) and extend their detected spectra up to the TeV energy range. At the same time Fermi-LAT provides essential information of interesting targets for ground based detectors, triggering observation proposals.

The current generation of IACTs is represented by **MAGIC**, **HESS**, and **VERITAS**.

- **MAGIC**: The Major Atmospheric Gamma-ray Imaging Cherenkov Telescopes (MAGIC) is located at the Roque de los Muchachos Observatory (La Palma, Spain) at 2200 m above sea level. It is a system composed by two IACTs of 17 m of diameter separated by 85 m. With a total mirror surface of 472 m² it represents the most sensitive IACT experiment between 30 and 300 GeV in the northern hemisphere, and also the biggest IACTs on Earth until HESS II was built.

Due to the huge size of the MAGIC telescopes, low energy events are observed and accurately reconstructed using stereo imaging, widening the low energy range of IACTs. This low energy threshold makes the experiment ideal for the detection of high redshift AGNs and γ -ray pulsar observations. MAGIC telescopes also have the fastest repositioning time among the IACTs, becoming the best candidate to observe the highest energy component of a GRB from Earth.

- **H.E.S.S.**: High Energy Stereoscopic System (H.E.S.S.) (named after Victor Hess) is formed by a system of 5 IACTs located in Khomas Highland (Namibia) at 1800 m above sea level, being the only major IACT experiment present in the southern hemisphere. The telescope layout is distributed as follows: four 12m of diameter telescopes form a perfect square of side equal to 120 m forming the classical HESS I system, and a 28m of diameter tele-

scope named HESS II sits in the center of the square. This system allows the observation of γ -rays with energies between 30 GeV to 100 TeV.

HESS II saw its first light on the 26th of July 2012, becoming the largest IACT on Earth, greatly improving H.E.S.S. sensitivity below 300 GeV. Its location in the southern hemisphere allows to observe key sky regions such as the galactic plane and galactic centre, the most interesting and populated areas of the sky in the TeV range. In 2009, H.E.S.S. was considered among the top 10 observatories worldwide ranked by their scientific impact by Nature [113].

- **VERITAS:** Very Energetic Radiation Imaging Telescope Array System (VERITAS) is built on Mount Hopkins (Arizona, USA) at 1268 m above sea level with an array of four 10 m of diameter IACTs placed in a diamond-like distribution with distances ranging between 80 to 120 m. This system of IACTs is able to observe from 50 GeV to 50 TeV, and is the most sensitive experiment in the northern hemisphere above 300 GeV.

2.3.2.1 Scientific results

Along with Fermi, γ -ray astronomy is mainly driven by IACTs. In the last decade, MAGIC, H.E.S.S. and VERITAS showed great performance with outstanding scientific results. As shown in Fig. 2.6, astronomy in the VHE regime is following a similar trend as other wavelengths underwent in the past in terms of number of detected sources.

Some scientific achievements of IACTs are briefly listed here:

- *Galactic:* The H.E.S.S. collaboration performed the Galactic Plane Survey [115], the first survey performed in the VHE range, detecting up to 14 previously unknown sources [116] and extending Fermi spectra located in that region to the TeV range. Another mayor discovery was the detection of the Crab Pulsar by MAGIC [25] and later on by VERITAS [117], constraining emission models and unveiling the origin of pulsed radiation.
- *Extragalactic:* The MAGIC collaboration observed fast variability from IC310 [51] in the VHE range constraining the size of the radiating region to be less

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

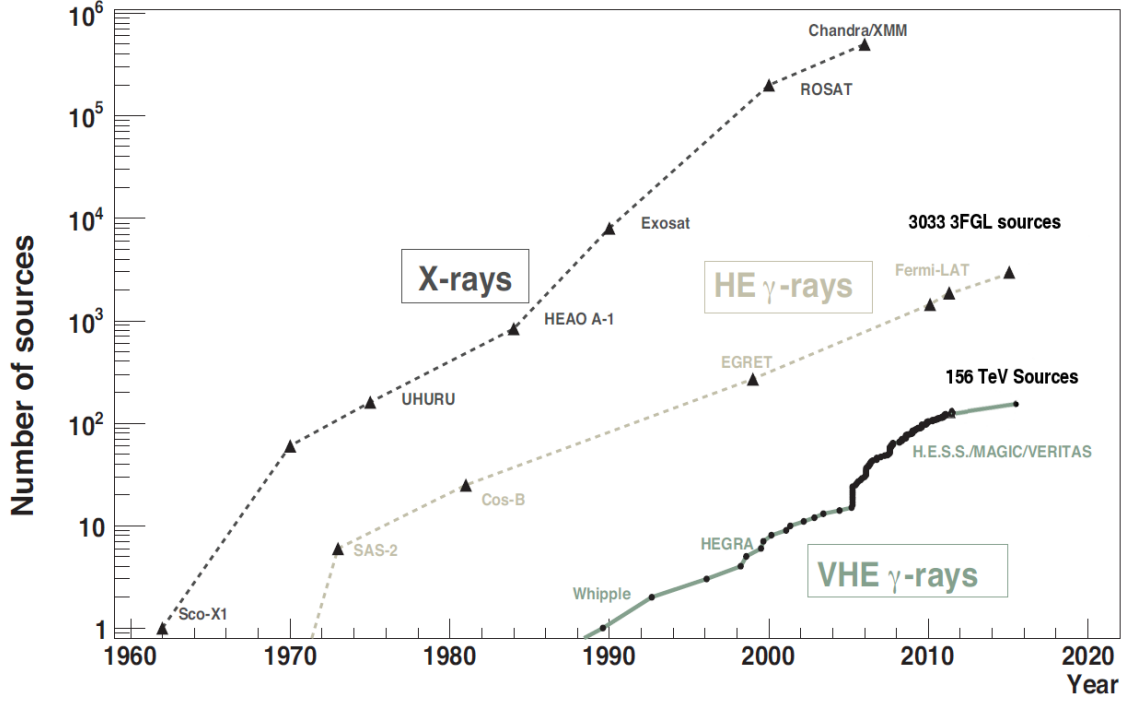


Figure 2.6: *Kifune* plot showing the number of detected sources for different astronomy energy bands as a function of the year. Updated version (22nd April 2015) based on [114].

than 20% of the gravitational radius of its central supermassive black hole, suggesting pulsar-like γ -ray emission processes. In addition, the detection of VHE gamma-ray spectra of high redshift AGNs performed by H.E.S.S. [52], MAGIC [53] [54] and VERITAS [55], [56] placed strong upper limits on the EBL density, indirectly measuring star formation rates along the history of the universe.

- *Fundamental physics:* IACTs also faced fundamental physics problems such as the search of Dark Matter, providing competitive constraints for certain super symmetric models of self annihilation WIMP particles through observations of the galactic center by H.E.S.S. [118] or dSph galaxies by H.E.S.S. [81], MAGIC [119] and VERITAS [120]. H.E.S.S. collaboration also performed measurements of the cosmic ray electron component [121] extending

the known electron-positron spectrum up to 5 TeV.

2.4 Future γ -ray observatories

There are currently two observatories in construction phase that will ensure γ -ray astronomy success during the next decades: HAWC and CTA, representing the new generation of Water Cherenkov detectors and IACTs respectively.

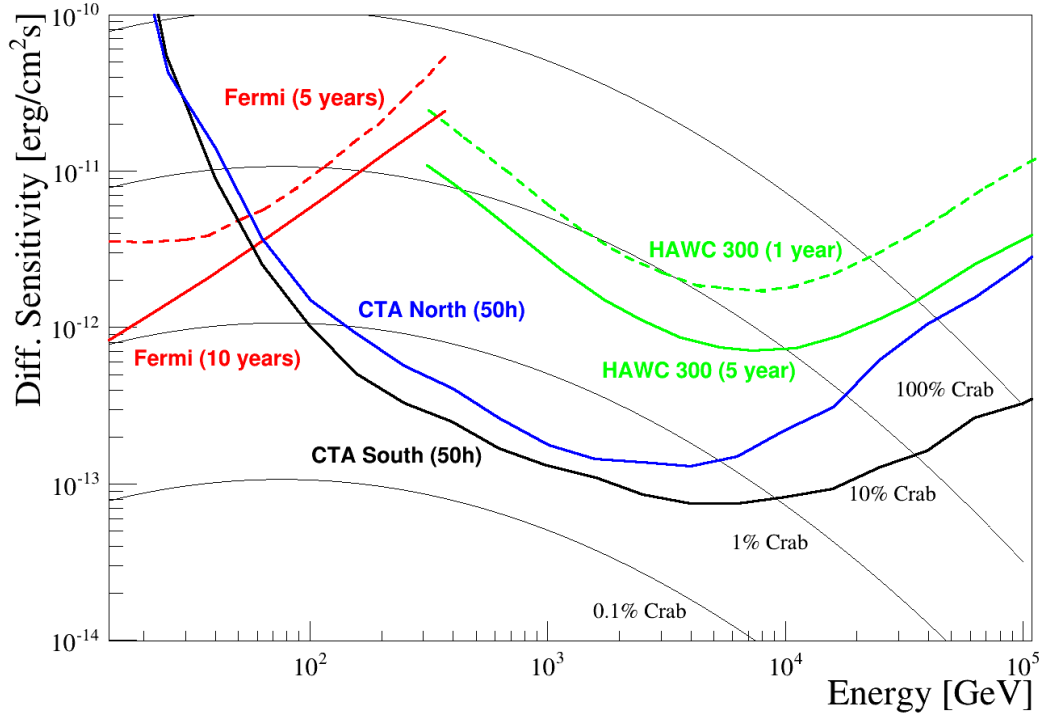


Figure 2.7: Minimum detectable flux of a point-like source for future VHE gamma-ray detectors. IACTs sensitivities correspond to 50h of observation while Fermi and HAWC corresponds to sky surveys through several years. CTA Northern and Southern hemisphere correspond to simulations of Tenerife 2N layout and Aar 2A layout observing at 20 degrees of zenith angle. Fermi and HAWC data obtained from [101], [122], [96].

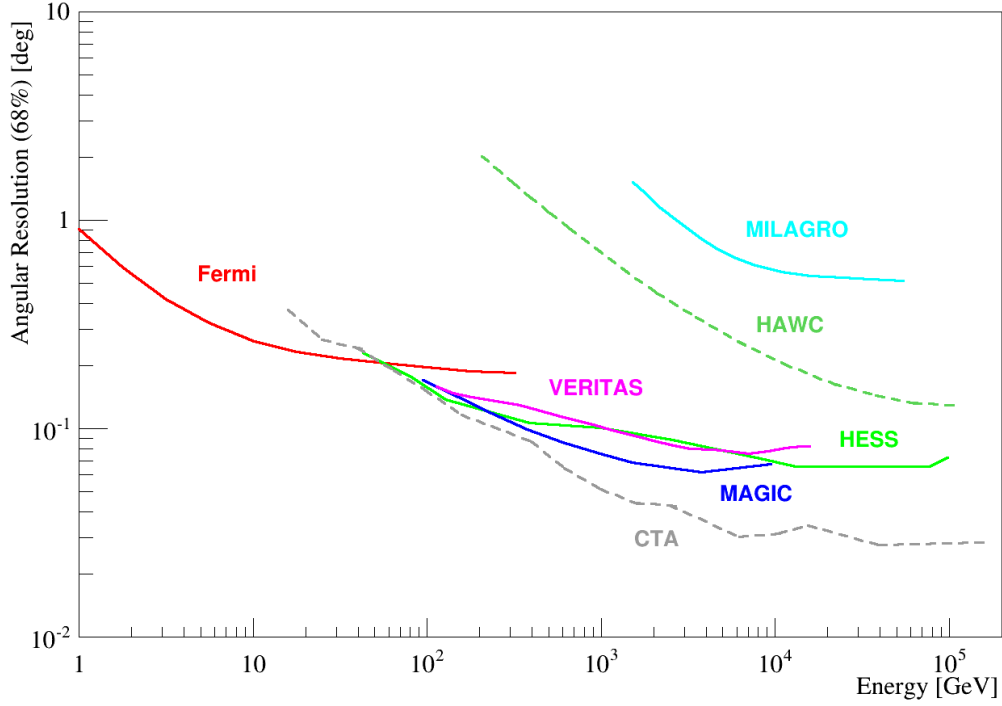


Figure 2.8: Angular resolution of present and future (dashed lines) VHE gamma-ray detectors. CTA angular resolution corresponds to 2A layout "AAR" site (Namibia). Data obtained from [101], [102], [94], [122], [96].

2.4.1 HAWC

High-Altitude Water Cherenkov (HAWC) observatory is located in Sierra Negra (Mexico) at 4100 meters of altitude, and will consist of an array of 300 individual large water Cherenkov detectors, as shown in Fig. 2.9b. The experiment observes γ -rays of energies between 100 GeV and 100 TeV with a wide field of view, being fully operational also during daytime, two-thirds of the sky are exposed every 24 hours. HAWC will extend Fermi-LAT survey in the northern hemisphere sky to the TeV range improving Milagro sensitivity over an order of magnitude, with improved angular resolution (see Figs. 2.7 and 2.8).

Developed by a collaboration of 15 North American and 12 Mexican institutions and funded by National Science Foundation (NSF), United States Depart-

ment of the Energy (DOE) and Mexican Consejo Nacional de Ciencia y Tecnología (CONACyT), HAWC was inaugurated in March 2015 and will perform an unbiased survey reaching already during the first year 50 mCrab sensitivity in the TeV range, similar to Fermi’s sensitivity in the GeV range.

As in other arrays of counters, in HAWC the primary particle direction is inferred using the relative time of arrival of the particles to the tanks (colored scale shown in Fig. 2.10). HAWC reconstructs also the shower core position (shown in Fig. 2.10 with a star marker). With this information and the measured intensity, the energy of the primary is estimated. Background suppression is performed searching for signal intensity asymmetries around the inferred core position. Isolated islands of particles measured far from the shower core are a hallmark of cosmic-ray events, as can be seen in Fig. 2.10b.

HAWC will explore a variety of scientific goals in the forthcoming years. The observatory possesses certain capabilities which allow it to be the perfect detector to study diffuse γ -ray emission and find TeV cosmic rays anisotropies in our galaxy, two probes that may unveil the origin of galactic cosmic rays. It will also explore extra-galactic objects, extending Fermi-LAT detected AGNs up to the TeV range. Its wide field of view and high duty factor allow the detection of extreme transients such as flaring states or even the brightest GRBs, information that may be of great value for other observatories without such capabilities, as the CTA observatory. These observations will also help to constrain the EBL and Inter-Galactic Magnetic (IGM), key parameters in cosmology.

2.4.2 CTA

As a result of the success of current IACT experiments and the improvement of the different technologies involved, the next generation of Cherenkov installations is under development. The Cherenkov Telescope Array (CTA), is a ground based observatory that will observe the VHE γ -ray sky from some tens of GeV up to a hundred TeV with unprecedented capabilities, improving current experiments sensitivity over an order of magnitude. Conceived by the CTA Consortium, a global initiative formed by more than 1200 people from 28 countries, CTA is currently being developed to attain ambitious scientific goals.

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

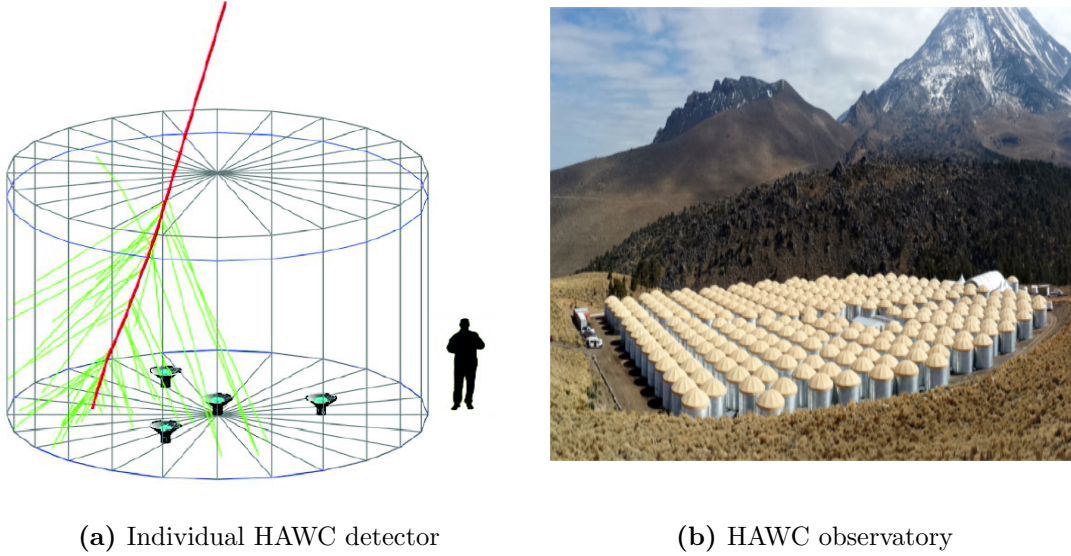


Figure 2.9: *Left:* Schematic view of an individual HAWC detector with an EAS induced particle (red) emitting Cherenkov photons (green) measured by PMTs located at the bottom of the tank. *Right:* Sky view of the future HAWC Observatory with a total of 300 tanks. Courtesy of the HAWC collaboration [96].

In order to achieve these goals the observatory will consist of two different sites, one in each hemisphere, consisting of 50 to 100 IACTs of 3 different sizes. This set up will provide full sky coverage with improved sensitivity alongside better angular and energy reconstruction. The CTA southern site is expected to be larger, composed of ~ 100 telescopes spread out over $\sim 6 \text{ Km}^2$, to take advantage of its latitude to observe the Galactic Plane and Galactic Center over the full energy range. The northern site is expected to complement it with a reduced number of telescopes, ~ 20 telescopes spread out over $\sim 1 \text{ km}^2$, more focused in low energy threshold studies such as AGNs, GRBs or starburst galaxies.

This project, along with HAWC and the extended operation of Fermi LAT shows the great boost γ -ray astronomy is undergoing in recent years. Fermi-LAT available data and HAWC future all sky survey will perfectly match with CTA point sensitivity, covering the whole VHE range with unprecedented detail, as shown in Fig. 2.7.

CTA will also be the first γ -ray open observatory ever constructed, allocating

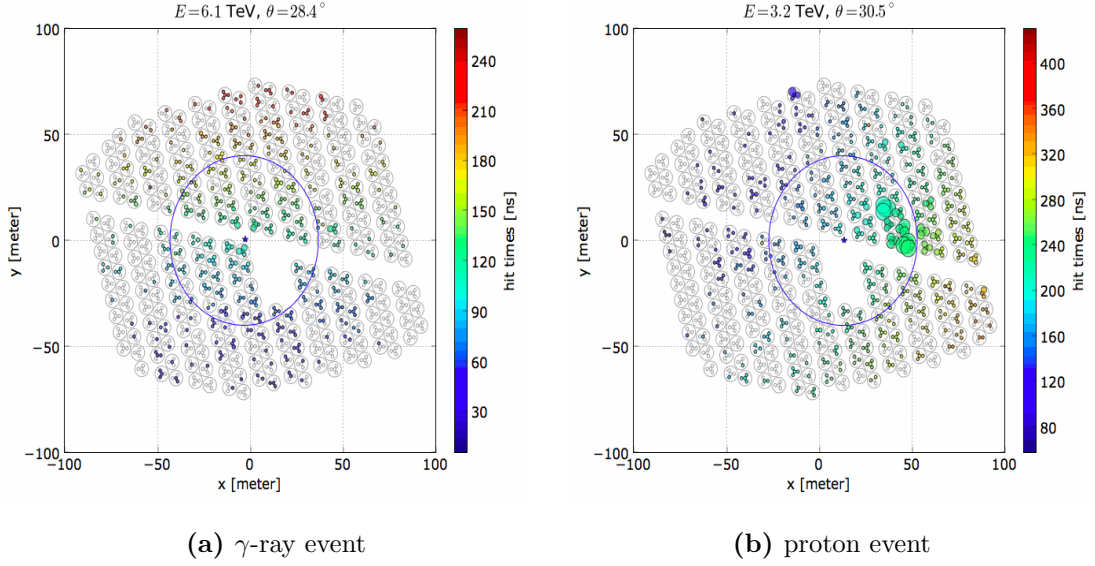


Figure 2.10: Simulated response of HAWC to γ -ray and proton events. Arrival time in ns is indicated by the color scale, while light intensity is denoted by the size of colored circles. Courtesy of the HAWC collaboration [96].

a significant fraction of observation time to external scientist proposals. This approach will bring the whole astronomical community closer to astro-particle physics, a field currently dominated by particle physicists, and boost scientific outcome. Consortium proprietary observation time will be devoted to calibration, technical measurements and several Key Science Project (KSP).

2.4.2.1 CTA Key Science Projects

CTA will make significant progress in all science cases withing the γ -ray astronomy, as the ones introduced in Sec. 1.3.1. In order to obtain the highest possible scientific outcome the CTA consortium proposed key observation projects gathering obvious scientific targets, of broad interest, to provide analyzed data and source catalogues that will trigger new proposals from the community. Proprietary time dedicated to KSPs will correspond to a very significant fraction of the commissioning phase and will gradually decrease afterwards. The KSPs currently proposed are the following:

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

- **Galactic Plane Survey:** Consists of a full survey along the galactic plane with deeper observations in the most promising regions between $-60^\circ < l < 60^\circ$. It may be considered the flagship project of the observatory with a number of important scientific goals. Hundreds of new sources should be detected in the broad energy range of CTA, mainly PWN and SNRs, along with new and unexpected γ -ray sources such as binaries, PeVatrons or transient phenomena in the Galaxy. The GPS KSP will be an undeniable legacy data set for follow-up observations and a long lasting value to the entire astronomical community.
- **Extragalactic Survey:** Nowadays, extragalactic source detections in the VHE range are often motivated by observations at lower wavebands, imposing a strong bias on the observed population. In addition, these sources are generally observed in transient flaring states. Therefore, an unbiased VHE extragalactic source catalog would be of great potential for the community. This project proposes a blind survey of 25% of the sky at high galactic latitudes $|b| > 10^\circ$, reaching a flux sensitivity for point like sources of ~ 5 mCrab, close to the weakest source seen by IACTs so far.
- **Deep observations of the Galactic Centre: probing astrophysics and dark matter:** The Galactic Centre may well be the most densely populated region in the γ -ray sky and at the same time the brightest source of DM annihilation (see Sec. 1.2.2). This KSP proposes deep observations of the galactic center with great potential of major discoveries, such as detecting a DM signature or revealing the nature of the gamma-ray source located at the dynamic centre of our galaxy.
- **Studies on Active Galactic Nuclei:** Active Galaxies amount to $\sim 40\%$ of the club of detected VHE sources and belong to the group of the most luminous objects in the universe. CTA has the potential to increase this category by a factor four and in particular to significantly increase the population of TeV sources at high redshifts. This measurements would help to understand acceleration processes taking place in the vicinities of black holes, the origin of extragalactic cosmic rays and fundamental questions such as determining

the EBL, the Inter-Galactic Magnetic field (IGMF) or the search of Lorentz invariance violation (LIV) and Axion-like particles (ALPs).

- **Observations of the Large Magellanic Cloud:** The LMC is the nearest star-forming galaxy, rich in γ -ray source candidates: dozens of SNRs, HII regions, bubbles and shells. Deep observations on the LMC will provide population studies on SNRs and PWNe, transport of cosmic rays along the galaxy and search for DM signatures.
- **Observations of transient phenomena:** Consists of follow-up observations, triggered by external alerts, of certain transient targets showing catastrophic events in the universe. Within the scope of the project there are known γ -ray emitters such as GRBs, flares from PWN, jet ejection events from X-ray binaries and novae, along with other exotic transients like high energy neutrinos or gravitational waves.

2.4.2.2 CTA performance

In order to achieve the ambitious goals described in Sec. 2.4.2.1, the CTA design target is to achieve certain goals, regarding different aspects of the observatory performance:

- **Sensitivity:** CTA is required to reach milli-Crab (mCrab) sensitivity in 50 hours of observation on a point-like source between ≈ 100 *GeV* to 10 *TeV*. This would improve the sensitivity of current generation of IACTs by a factor 10 inside the CTA core energy range. This unprecedented sensitivity will allow the study of fainter objects, already observed by Fermi-LAT, discover new source populations of even fainter ones and reduce the selection bias of the different γ -ray source types.
- **Angular resolution:** In order to resolve extended sources such as SNRs, CTA should be able to reach angular resolutions below 2 arc minutes for energies above 1 *TeV*, the best resolution ever achieved in this range of energies, improving by a factor 3 the usual values attained in current instruments.

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

- **Energy range:** The observable energy range should extend well below 100 GeV and up to more than 200 TeV, with improved energy resolution with respect to current IACTs. CTA will be able to observe more than 3 orders of magnitude in energy, crucial for discerning between different emission mechanisms scenarios in AGNs or distinguish different hypothesis of gamma ray production in SNRs.
- **Temporal resolution:** CTA large effective area and improved angular resolution will open a new window to a wide variety of transient phenomena. Current IACTs have already measured short time-scale variations of few minutes in the most rapid flux variations ever observed in AGNs, constraining the emitting region size [51]. CTA will be able to resolve sub-minute time scales, improving constraints on AGN emitting region and significantly increase the chances of detecting extreme transients such as the very high energy component of GRBs, never observed from ground.

These performance goals represent an enormous technological challenge. We can divide the CTA energy range in three regions, each one corresponding to a different limiting factor:

The low energy range ($E < 100$ GeV): γ -rays below 100 GeV produce showers with low Cherenkov photon density, typically concentrated in a ring of radius $R \approx 120m$. Individual IACTs need large reflecting areas in order to integrate enough photons to discern between Cherenkov photons and Night-Sky Background Light (NSBL). In addition, the overwhelming amount of low energy hadronic showers produce images similar to the ones from γ -rays, turning background rejection into a difficult task. In this energy range sensitivity is limited by background rejection and signal to noise ratio.

CTA will push the low energy threshold down to some tens of GeV using a system of LST, each one equipped with a large mirror area (23 m of diameter) to collect enough Cherenkov photons from low energy showers. Since at low energies event rates are high, large collection areas are not needed, and a few number (≈ 4) of LSTs placed at short distances (≈ 100 m) will be able to provide enough effective area ($\sim 10^4$ m²) and decent background rejection.

The medium energy range ($100 \text{ GeV} < E < 10 \text{ TeV}$): Corresponds to the core energy range of CTA. This region is well understood from the experience gained by the current generation of instruments. Cascades generate higher Cherenkov photon density, consequently reflectors with very large areas are not needed anymore. Images, composed of more pixels, reflect more features of the shower development through the atmosphere, helping to discern between hadron and γ -ray events. Sensitivity is mainly limited by the effective area and the quality of stereo reconstruction.

To maximize the core energy of CTA, a system of medium-sized telescopes (MST) of about 12m diameter will be spread over an area of $\approx 5 \cdot 10^5 \text{ m}^2$. Each cascade should be stereoscopically imaged by as many telescopes as possible in order to improve the quality of shower reconstruction, so telescope separation should range between $100 - 150\text{m}$. For the first time, the area covered by the IACT system will be significantly larger than the Cherenkov light pool size, and therefore most of the events in this energy range will have their impact point reconstructed inside the array, improving the quality of the reconstruction.

The high energy range ($E > 10 \text{ TeV}$): At these energies, cascades generate a huge amount of Cherenkov photons, becoming observable at larger distances. The images show clear distinctive features in the EAS development between γ -ray and cosmic rays, turning background rejection into a simple task. As a consequence, sensitivity is only limited by the total effective area telescopes are able to cover, only constrained by the requirement of stereo shower reconstruction.

To maximize effective area, CTA will use small-sized telescopes (SST) of small reflecting area ($\sim 10 \text{ m}^2$) separated by a wider distance than other telescope types. This distance is yet to be optimized, but it will range between 200 and 400 m, depending on the final model chosen.

This thesis work is devoted to the evaluation of CTA performance through detailed MC simulations and the analysis of its future scientific impact in γ -ray astrophysics. Chapter 3 describes the analysis performed and shows the impact on the sensitivity of different elements involved in the design phase of CTA, studying the effect of parameters associated with the construction site (altitude of construc-

2. PAST, PRESENT & FUTURE OF γ -RAY DETECTORS

tion, geomagnetic field intensity or Night-Sky Background (NSB) level) or array layout: different telescope types, spacing and distribution. Chapter 5 is devoted to explore MLA within γ -ray astronomy, evaluating their current applications and probe their implementation of new purposes. In chapter 4 CTA capabilities are evaluated in different physics cases, estimating the final scientific impact of the observatory.

Chapter 3

Sensitivity studies for the CTA

This chapter is the central part of this work, and gathers all the sensitivity studies performed through MC simulations and analysis for the CTA. First, section 3.1.1 introduces the IACT technique, describing the main characteristics of these telescopes and how CTA is planned to improve the performance of the current generation of γ -ray detectors. Then, section 3.2 describes the different software packages involved in the CTA MC simulations, from the development of the EAS to the response of each telescope measuring the Cherenkov radiation. Section 3.3 describes the different tools used for the analysis of the CTA MC productions, both the software used in this work and the existing alternatives. An overview of the different large-scale MC productions performed by the CTA collaboration together with their configuration is performed in section 3.4. The two last sections present the main results of this work: Section 3.5 analyzes CTA performance in detail, showing the different Instrument Respond Functions (IRFs) describing the observatory capabilities and how these are affected by the MST types used in the telescope layout. Section 3.6 concludes this chapter by evaluating the different construction sites proposed for the CTA-N observatory, studying the effect on performance of the location related parameters such as the altitude or the Geo-magnetic field.

3.1 Imaging Atmospheric Cherenkov Telescopes

IACTs currently provide the best sensitivity among the different detection techniques in the VHE range. As previously introduced in section 2.2.2, Cherenkov photons from EASs are collected with wide mirror surfaces and focused on a camera, where these photons are measured creating images of the shower development. They operate in similar fashion as optical telescopes, although the short timescale of the EAS Cherenkov flashes and the overwhelming cosmic ray background impose certain key differences, mainly regarding the camera nature and data analysis needed. The technique is considered rather new, as they have been operating for less than half a century, but the low cost of the experiments compared to space detectors, and the benefits they provide in terms of effective area and sensitivity pushed this technique forward, specially in the last decades.

A general overview of the IACT technique and data analysis will be done in the following sections. For a deeper description, the reader is encouraged to glance over more specific experiment overviews, such as [123] of the MAGIC telescope.

3.1.1 Technical description

As already highlighted in section 2.2.2, IACTs indirectly measure γ -rays by observing the Cherenkov flashes they produce as they collide with the Earth's atmosphere. These are observed with systems similar to classical telescopes, focusing Cherenkov photons with a large reflecting area into a camera. But it is impossible to understand the design of IACTs, very different than classical telescopes, before deeply understanding the technical limitations they need to confront. The reader is encouraged to take a look at the appendix A for a better understanding of Cherenkov emission and the development of the EASs.

The greatest technical limitation of IACTs is the short time-scale of the Cherenkov flashes. Increasing with energy, they last from few ns up to some tens of ns. As a consequence, camera electronics are restricted to work at very high frequencies (unlike optical telescopes). In addition, fast pixel information needs to be retrieved to discern between Cherenkov flashes and other types of radiation.

The distinctive features of the IACT systems is constrained by the different

sources of background affecting them. These can be divided into 2 groups, depending on their nature. The first group of background sources listed here accounts for the different particles impinging the atmosphere generating EASs:

- *Hadronic induced EASs*: Cosmic Ray (CR) hadronic showers are the main source of background for the IACTs. Due to the similar development of γ -ray and hadron induced EASs, background cannot be suppressed by direct methods, requiring a relatively complex data analysis that exploits geometrical differences between hadronic and electromagnetic showers development, described in A.2. Hadron induced EASs account for the 99.9% of observed background and is efficiently suppressed by parameterizing image parameters and using multi-variate analysis (MVA) methods to discriminate between γ -ray and hadron induced cascades.
- *Electron induced EASs*: Within the cosmic ray radiation there are high energy electrons also generating EASs. These showers are purely electromagnetic, and differentiating between γ -ray and e^- induced cascades becomes an impossible task for the current generation of experiments. As shown in Fig. A.6, in the VHE range a very small fraction of the total background corresponds to electrons. Although their contribution to the total CR spectrum is small, their impact on IACTs sensitivity is significant at the energy ranges where hadrons are most efficiently suppressed, being the dominant source of background for the CTA around 1 TeV.
- *Diffuse γ -ray background*: A third source of EASs, in this case irreducible, is the diffuse γ -ray background. It consists on two components of galactic and extragalactic origin. Models seem to be in conflict with the observed diffuse γ -ray background, so the origin of this emission is still unknown. Their contribution is significant in certain regions, for instance the galactic center, where it needs to be fitted and subtracted.
- *Muons*: They are produced in hadronic showers, reaching the ground before decaying into an electron and a neutrino. Although they do not generate electromagnetic showers, they do emit Cherenkov photons along their trajectories. This Cherenkov light produces images in the IACT cameras depending

3. SENSITIVITY STUDIES FOR THE CTA

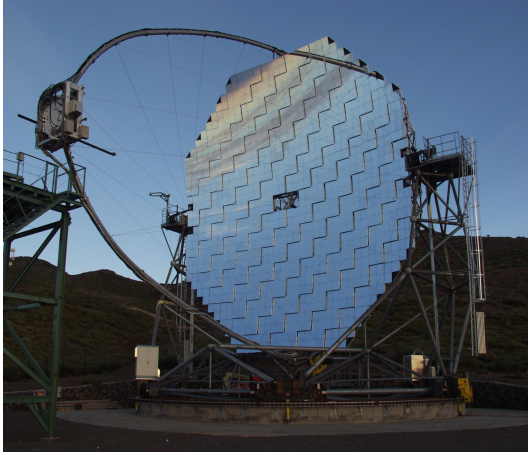
on the muon trajectory, with respect to the telescope position. Muons passing near the telescope create *muon rings* images, easily suppressed, while farther away muons may mimic low energy γ -ray events becoming another source of background (heavily suppressed by stereoscopic imaging).

The second group of background sources do not correspond to developed extended showers or Cherenkov emitting particles, but also affect to the IACT performance in various ways:

- *Night-Sky Background Light (NSBL)*: Corresponds to the visible light coming from the diffuse scattered light from stars, clouds, Moon light or artificial sources such as cars or cities in the vicinity of the observatory. These photons enter the telescope optics and mimic low energy events, creating accidental triggers easily suppressed in the analysis, but affecting the data acquisition.
- *Electronic noise*: Camera electronics show certain levels of noise that need to be taken into account. High electronic noise levels may create accidental triggers reproducing the effect of NSBL. This noise needs to be characterized using daily dedicated observations to minimize its effect.

As a consequence of these background components and the EAS nature, IACTs are designed to collect Cherenkov photons within very short time-lapses, gathering as much geometrical information of the shower development as possible to discern between γ -rays and other Cherenkov emitters. IACTs are composed of 3 main systems:

Mirror surface: In order to collect as many Cherenkov photons as possible, a wide reflecting surface of segmented mirrors is used. The mirror segments are aligned by independent actuators built into each support cell, reducing the effect of structure deformations. Classically, the geometric shape of the reflecting surface is parabolic, locating the camera in the focus, as shown in Fig. 3.1a for the case of the MAGIC telescope. This way Cherenkov photons emitted at a certain altitude arrive at the same time to the camera, reducing the time spread of the signal (reducing the integrated NSB). The arrival time of Cherenkov photons also provides geometric information of the shower development (improving background



(a) MAGIC Telescope.



(b) SST 2m prototype.

Figure 3.1: *Left:* One of MAGIC telescopes, with a segmented mirror surface of 17 m in diameter and PMTs camera located in the focal plane. *Right:* Schwarzschild-Couder SST prototype, with a 4 m in diameter of a segmented primary mirror with a digital Silicon Photo-Multiplier (SiPM)s camera located near the secondary mirror, built in Catania by the Astrofisica con Specchi a Tecnologia Replicante Italiana (ASTRI).

rejection). New designs show promising results using Schwarzschild-Couder [124] double mirror optics. This system allows a wider FoV and reduces comatic aberrations, improving imaging resolution. An example of this double mirror setup can be seen in Fig. 3.1b.

Camera: It is the most distinctive and crucial element of the IACTs, and its performance conditions the overall sensitivity of the instrument. Due to the short time-scale of Cherenkov flashes, required data acquisition rates and the large plate scale, classical cameras of higher Quantum Efficiency (QE) cannot be used (such as CCD cameras), so Photo-Multiplier Tube (PMT) based pixels are used instead, with typical QE values of $\sim 35\%$. PMTs amplify the signal proportionally to the number of incident photons, transforming them into analog signals, allowing ultra fast acquisition rates. Pixels may use additional light collectors, such as Winston cones or lenses, in order to increase collected photons entering each individual

PMT.

Data readout: The electric signals generated by the PMTs are handled by the readout system to ultimately digitize and store their output by the analog to digital converter (ADC) at very high frequencies. The PMT signals are sent to the trigger system and buffered at the same time. The trigger evaluates, using information from other pixels, if the signal is worth to be stored or if its likely to be background. If the signal triggers in several IACTs (in stereoscopic systems) buffered data is digitized and stored.

3.1.2 Stereoscopic imaging principle

Since HEGRA developed the stereoscopic imaging in the end of the XX century, all competitive experiments incorporated this technique. Individual IACTs provide independent images of the EASs, but this information is insufficient to accurately reconstruct parameters such as direction, core position and altitude of the shower maximum, key parameters to understand the geometry of EASs.

The core energies of the CTA will improve the sensitivity of the current generation of IACTs by an order of magnitude, mainly by increasing the amount of telescopes observing each event (*multiplicity*). As shown in Fig. 3.2, shower images are parametrized (fitted to an ellipse using Hillas parameters [127]) and the information they carry is combined to reconstruct the impinging particle's properties. Each individual image observing the same EAS is carrying information of a different projection of the shower where Cherenkov emission was generated. Having several points of view of the same event significantly improves every step of the primary reconstruction:

- Direction reconstruction: As shown in Fig. 3.2, each camera image provides a projection of the direction of the shower. The contribution of more projections from the same event will lead to a better determination of the *source position*, *core position* and *maximum height*, and therefore better estimation of the *impact parameters*.
- Energy reconstruction: In this case, the most important source of information is the amount of Cherenkov light detected from the shower at a given

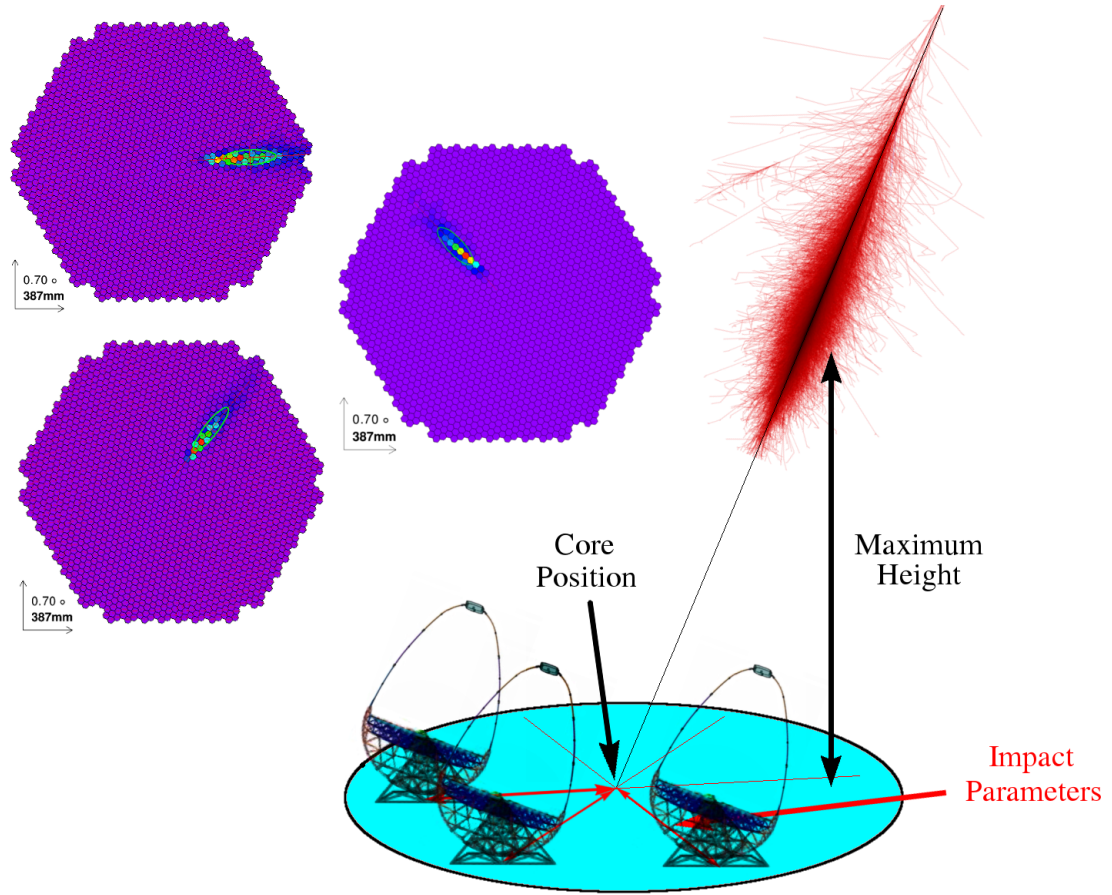


Figure 3.2: Schematics of the stereoscopic imaging principle used by IACTs. Individual telescopes produce pixelated images of a projection of the development of the shower. Images are parameterized and combined to reconstruct the original direction of the incident particle defining its *core position*, the *maximum height* together with the *impact parameter* of each IACT. Red traces show ultra-luminic particles produced within a EAS produced by a 1 TeV impinging γ -ray, generated with CORSIKA [125], courtesy of F. Schmidt [126]. Camera images show the cleaned image and parameterized ellipse of a 1 TeV γ -ray seen by 3 different CTA large size telescopes.

3. SENSITIVITY STUDIES FOR THE CTA

distance. Together with the improved *impact parameter* determination, having more telescopes allow the detection of a higher percentage of the emitter Cherenkov light, leading to an enhanced energy reconstruction.

- Gamma-hadron separation: In the low energy range, the projection of hadronic and electromagnetic cascades may look very similar, and differences may be unveiled only from specific directions. Having more telescopes improves the chances of catching those differences. Also, the detection of muons or Π_0 sub-showers generated from hadronic EASs is a signature to label an event as background.

In addition, using several IACTs enables the option of stereo triggering, only storing events detected by several telescopes. This restriction reduces the amount of γ -rays observed (effective area of the instrument), but overwhelmingly rejects background generated triggers, mainly produced by accidentals or muons, significantly improving trigger efficiency and therefore sensitivity.

3.1.3 CTA telescopes

The CTA represents the future of VHE γ -ray detectors. The CTA Consortium is currently developing the next generation of IACTs, implementing new optic system approaches, new cameras with alternative PMTs and improved readout systems. As briefly introduced in Sec. 2.4.2, CTA needs different telescope types to fulfill the ambitious scientific requirements its designed for. In order to have such a wide energy coverage, from some tens of GeV to a few hundreds of TeV, different telescope sizes were proposed, each improving CTA sensitivity at different energy ranges. Irrespective of the technical implementation details, the initially proposed telescope types come in three sizes: The large-sized telescopes (LST), the medium-sized telescopes (MST) and the small-sized telescopes (SST).

LST: With a 23 m diameter reflector, it is the largest CTA telescope, designed to detect the lowest energy EASs, dominating CTA sensitivity below ~ 100 GeV. The proposed camera will be composed by ~ 1800 PMTs intended to work at high frequencies in order to trigger as many low energy events as possible (with

acquisition rates about ~ 10 kHz in stereo). The mirror structure will be parabolic to keep the isochronicity of the optics, narrowing the trace integration time window and therefore reducing the NSB contribution. With a 28m focal length it is expected to have a FoV of about 4.5° . The baseline design is shown in Fig. 3.3. The requirement of fast re-positioning time (for rapid GRB follow-up) imposes the choice of a light weight structure similar to the one built in MAGIC. Few number of these telescopes (probably ~ 4 at each site) placed with a moderate separation (~ 100 m) are expected to be built in each Hemisphere. The construction of the first prototype is currently on-going in *El Roque de los Muchachos* (La Palma, Spain) close to the MAGIC telescopes.

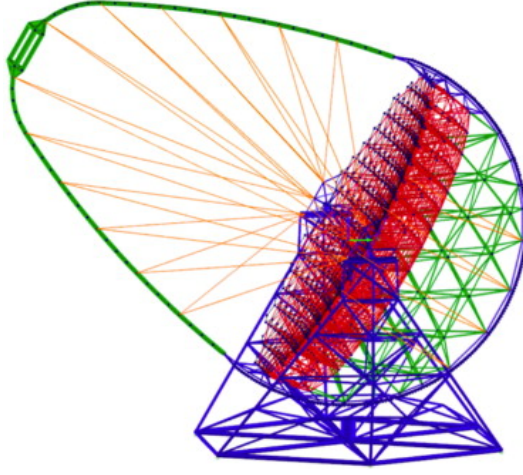


Figure 3.3: Baseline design of the LST, with a 23 m of diameter parabolic mirror surface.

MST: As the energy of the impinging primary particle increases, the amount of Cherenkov light emitted by EASs augments, and large telescope reflecting areas are not needed anymore. The MSTs are designed to increase the effective area and the number of telescopes simultaneously observing each event, dominating the sensitivity from ~ 100 GeV up to ~ 5 TeV. Two designs have been proposed and developed: The Davies Cotton (DC)-MST (or simply referred as MST indistinctly), a 12 m diameter single mirror IACT with a modified Davies-Cotton optics (see Fig. 3.4a), improving its off-axis performance by only adding a nearly

3. SENSITIVITY STUDIES FOR THE CTA

negligible time spread to the Cherenkov photons time of arrival. With a focal length of 16 m it will have a FoV of $\sim 8^\circ$. The MST camera will be similar to the one in the LST, using ~ 1800 PMTs. The second design is the Schwarzschild-Couder medium-sized telescopes (SC-MST) the first ever designed double mirror IACT (see Fig. 3.4b). This telescope features a novel two-mirror optical design that fully corrects spherical and comatic aberrations while providing a large field of view (FoV). After reflecting from its 9.7m diameter primary mirror the de-magnification of the shower images performed by the secondary mirror provides a fine plate-scale that allows for a highly pixelized focal plane instrumentation. The current design allows a camera with ~ 11000 Silicon photomultiplier (SiPM) pixels with a FoV of $\sim 8^\circ$, capable of recording the shower development with unprecedented image resolution. The MST will populate the central part of the CTA layouts: ~ 15 are expected to be built in the Northern Hemisphere, increasing that number up to ~ 25 in the Southern site with a spacing optimizing the best trade-off between event quantity and reconstruction quality. An extension of SC-MSTs is expected to be constructed probably surrounding the central MSTs.

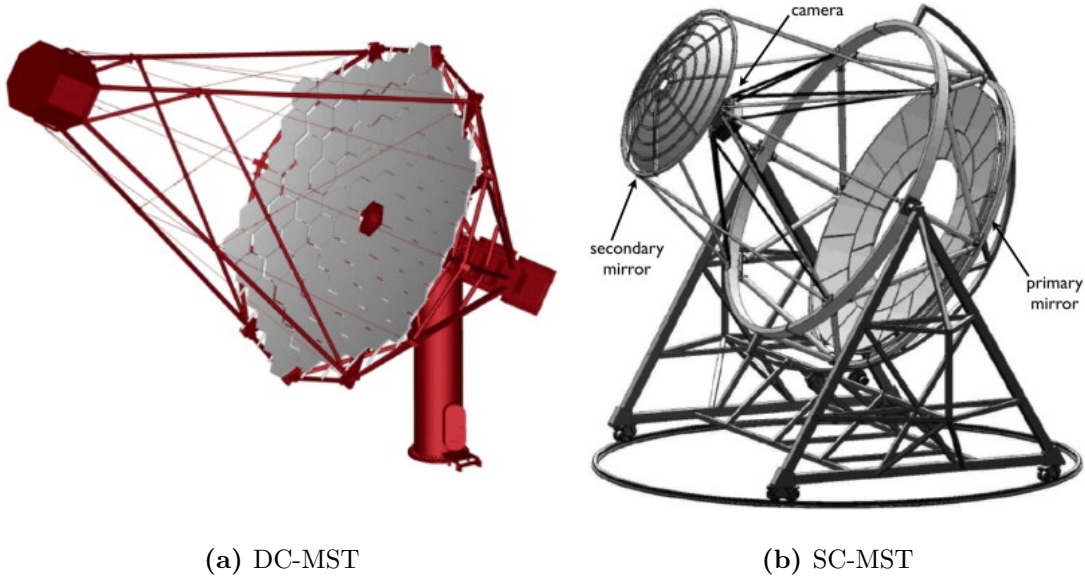


Figure 3.4: Proposed MSTs by the CTA Consortium: *Left:* Baseline design of the DC-MST, with 12 m diameter modified Davies-Cotton optics. *Right:* Baseline design of the SC-MST, with a 10m diameter Schwarzschild-Couder double mirror optical systems.

SST: Above a certain energy, small sized IACTs are able to observe the Cherenkov light generated by the EASs, and reconstruct the primary particle properties efficiently, taking into account that at these energies hadronic background is negligible. They are required to provide γ -ray collection areas of several km^2 , dominating the CTA sensitivity above $\sim 5 TeV$. Initially, 7m-SSTs were being considered but budget constraints ended up with several 4m prototypes, currently being developed (shown in Fig. 3.5): The DC-SST mounts a 4 m diameter Davies Cotton optics with 5.6 m of focal length and a FoV of 9° projected in a camera with ~ 1300 hexagonal SiPM. The second proposed small-sized telescope is the Compact High Energy Camera (CHEC) Schwarzschild-Couder small-sized telescopes (SC-SST), mounting a 4 m in diameter double mirror with a camera of 2048 multi-anode photomultipliers (MAPM) pixels and a FoV of about 9° . An independent design of a double mirrored telescope is the ASTRI 4 m SC-SST, with a camera of 2048 SiPM pixels with a FoV of 9.6° with an already working prototype pictured in Fig. 3.1b. Many (~ 70) of these telescopes will cover a very wide area ($\sim 5 km^2$) in the Southern Hemisphere site.

3.2 Monte Carlo simulation tools

The Monte Carlo simulations used in the IACT technique comprise a group of software tools that study the development of EASs in the atmosphere and the IACTs response. Starting at the composition, energy and direction of the primary particle, the cascade is simulated taking into account interactions and decays of generated nuclei, hadrons, electrons and muons, and their corresponding Cherenkov emitted photons. After these photons reach the ground, telescopes optics and ray traces are computed along with individual pixel responses.

The Monte Carlo (MC) simulations have a leading role in the design phase of CTA. Accurate performance estimators can be calculated through detailed MC simulations and analysis, testing different key parameters for the design of the observatory, such as the construction site altitude, the geomagnetic field effect or differences between telescope layouts and types. Ultimately, the objective of the CTA MC working group is to find the most cost-effective telescope layout, which maximizes the overall performance for a given cost, fulfilling all requirements

3. SENSITIVITY STUDIES FOR THE CTA

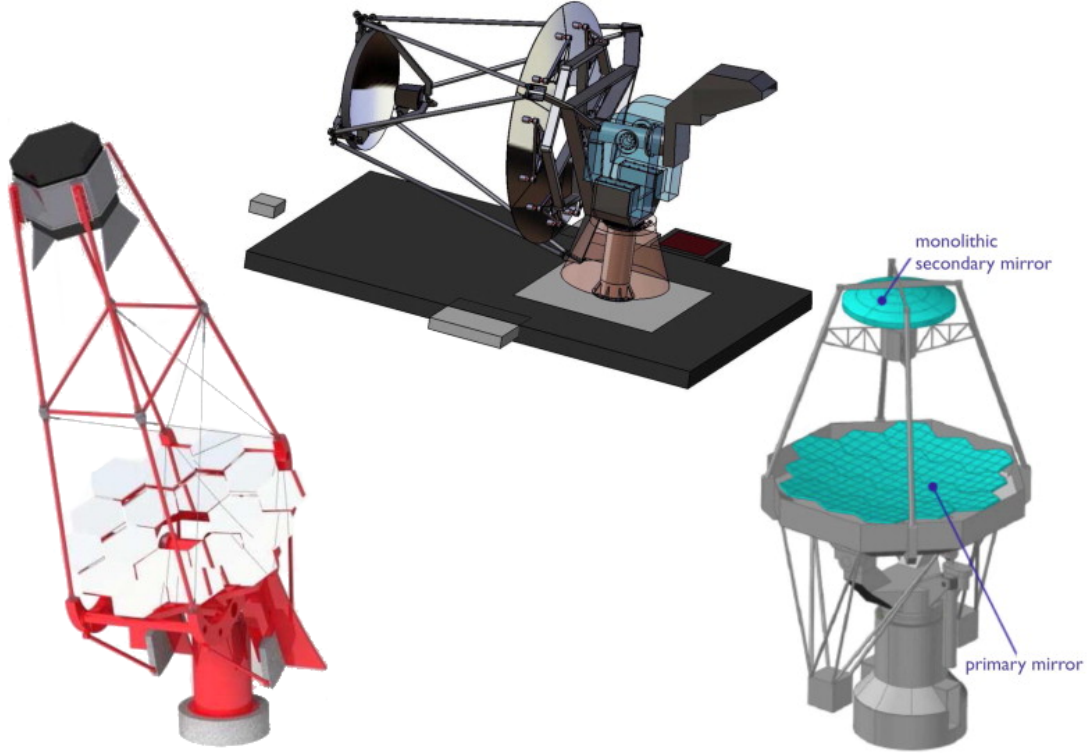


Figure 3.5: Different 4m-SST types currently being developed by the CTA Consortium: *Left:* Baseline design of the Davies-Cotton DC-SST. *Top and Right:* Baseline design of the double mirror CHEC SC-SST and ASTRI SC-SST respectively.

imposed by the Consortium.

These simulations have been of great importance to the IACT technique. Unlike other experiments where the direct observation of γ -rays allow the instruments to perform a straightforward calibration in test benches (as with scintillators, Compton scattering detectors, etc...), the IACT analysis needs to rely on MC simulations in order to reconstruct the primary particle properties and assess the telescope performance.

The main uses of MC simulated data in the current generation of IACTs analysis are the following:

- **Gamma-hadron separation:** One of the greatest disadvantages of the

IACT technique is the similar nature of γ -ray and cosmic ray induced EASs. The most efficient way to discern between different primary particles is to compare the image parameters observed by the different telescopes with the ones generated through MC simulations. Using MVA methods is possible to discriminate, up to a certain degree, between γ -ray and hadronic showers.

- **Direction reconstruction:** In order to compare the reconstructed direction of the primary particle with respect to its original direction, MC simulated γ -rays are needed. They are used to calculate the IACT angular resolution, and also to train more sophisticated methods to calculate the arrival direction, such as Look-Up Table (LUT)s or the *DISP* method [102, 128].
- **Energy reconstruction:** To estimate the energy of the primary particle, similarly as in the gamma-hadron separation case, shower parameters are compared to the ones obtained with simulated events using MVA methods. Taking into account this method is not perfect, an error is added to the energy calculation, creating a bias in the experimental measurements. For this reason, this energy is corrected by unfolding methods [129, 130], which also need MC generated data.
- **Performance studies:** Overall performance indicators as the effective area, background rates or the angular and energy resolution of IACTs are also calculated using MC simulations in order to estimate the sensitivity of the instrument. They also provide a test bench for different observation techniques, image cleaning methods or alternative signal extraction approaches.

CTA Monte Carlo simulations consist of 2 different stages. The shower simulation is performed using CORSIKA, briefly explained in Sec. 3.2.1. The CTA telescopes response is emulated using `sim_telarray`, and will be also introduced in Sec. 3.2.2.

Once simulations are performed, the MC production needs to be analyzed following similar methods to those applied for real data from the current IACT experiments. In this work MAGIC Analysis and Reconstruction Software (MARS) software package [131] based tools are used, detailed in Sec. 3.3.3.

3. SENSITIVITY STUDIES FOR THE CTA

3.2.1 CORSIKA

The simulation of all γ -ray, hadronic and leptonic showers of the CTA MC were computed using the COsmic Ray SIMulations for KAScade (CORSIKA) ([125], [132]) software package. Originally developed for the KASCADE experiment [133] at Karlsruhe (Germany), this software has evolved by the input of many users from different experimental fields. It was chosen taking into account the variety of interaction models available and because it has been widely tested by many different experiments.

CORSIKA is a detailed Monte Carlo simulation program of EASs. It allows the use of many different primary particles (protons, γ -rays, nuclei up to Fe, e^- , e^+ , muons, etc...) which are tracked through the atmosphere along with the subsequent interactions with the air nuclei, decays and secondary particle production. As an example, Fig. 3.6 shows 3 different showers generated with CORSIKA, each corresponding to a different impinging particle of an energy of $E = 1 \text{ TeV}$. It also allows to configure many physical parameters affecting the showers development and observed Cherenkov photons such as the observation altitude, the atmospheric density profile and the local geomagnetic field intensity, all of them crucial parameters for appraising the proposed CTA construction sites.

In order to test different telescope arrays and estimate their performance, a super layout of telescopes is designed, composed by the superposition of reasonable positions covering a total area of $\sim 7 \times 10^6 \text{ m}^2$. Several MC mass productions were performed, each having different scientific objectives, with different physical properties and simulated telescope number, position and types. Fig. 3.12 pictures the telescope layout used in the First MC Production (Prod-1) and Fig. 3.38 shows the one used for the Second MC Production (Prod-2) (these productions will be introduced in Sec. 3.4). Using these macro-layouts, CORSIKA's output files only need to store the Cherenkov photons at ground level in the vicinity of the telescope positions, considerably decreasing the output size.

To properly estimate the CTA performance, a huge number of cascades needs to be generated (in the order of $\sim 10^{11}$ cosmic rays and $\sim 10^{10}$ γ -rays), demanding a huge amount of computation time. Simulations were performed by the MC working group using the CTA computing Grid (part of the LHC Computing Grid

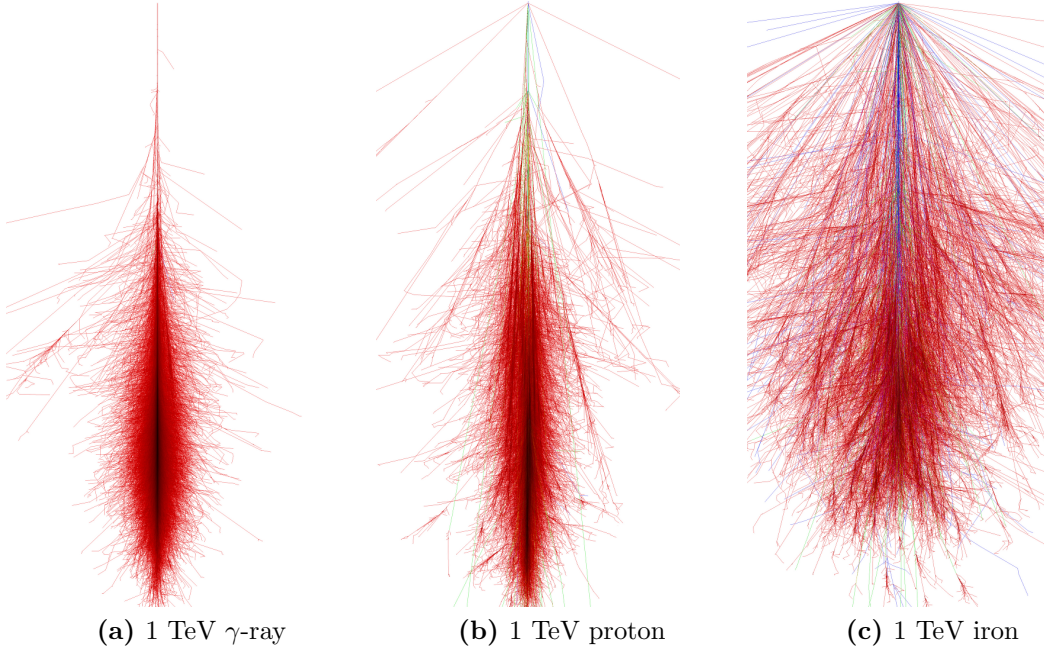


Figure 3.6: Horizontal projection of the shower development of 3 different impinging cosmic particles (or nuclei) of 1 TeV, simulated using CORSIKA. The color scale of the particles traces is the following: *Red*: electrons, positrons or γ -rays. *Green*: muons. *blue*: hadrons. Images courtesy of F. Schmidt [126].

(LCG) [134]). CORSIKA version 6.0 and up was used.

3.2.2 `sim_telarray`

The second step of the CTA MC production is to simulate the response of individual telescopes to the CORSIKA generated Cherenkov photons. This simulation is performed using `sim_telarray` [135], software developed mainly by K. Bernlöhr for the High Energy Gamma Ray Astronomy (HEGRA) telescope and then adapted and used by the H.E.S.S. experiment. Its efficiency, inherent flexibility and configurable settings allow the integration of new telescope types, sizes, optics systems, and new cameras or reflectors designs.

`sim_telarray` simulates telescope response with accurate detail, simulating the ray-tracing of different reflector optics and setups, trigger conditions for different discriminator thresholds and array triggers, adds the NSBL and electronic noise

3. SENSITIVITY STUDIES FOR THE CTA

and finally computes each PMT response individually emulating the digitization.

In order to assess CTA telescope capabilities, `sim_telarray` allows the configuration of each telescope independently and measures proposed instruments response up to an incredible degree of precision. Here we enumerate the different steps performed in the simulation and the different configurable parameters:

- Reflectors and ray tracing: `sim_telarray` performs the ray tracing and allows highly configurable reflecting systems, including individual segmented reflectors layouts (of different shapes or focal lengths), dual mirror optics or even masts and camera shadowing. Fig. 3.7 shows the degree of detail ray tracing simulation is able to reproduce the IACTs optical systems. Mirror reflectivity curves are used to account for the wavelength dependence of the different mirrors aluminizations and coatings. If required, this software is also able to calculate the optical PSF in the camera plane of a point-like source, taking into account deviations of mirror alignment. Ultimately, the distribution of Cherenkov photons, arrival direction and times are calculated at the focal plane.
- Camera: The camera is also defined by individual pixels with configurable position, size, separation, type and shape (circular, square, hexagonal). Funnel are also evaluated independently, taking into account their reflectivity and angular acceptance. The pixel type fixes the QE used, the single photo-electron response and the signals shape.
- Trigger and digitization: Different trigger schemes are also tested in order to compare their efficiency and sensitivity. Majority, analog sum, and digital sum triggers are implemented. Each scheme uses a different signal digitization and allows the configuration of independent trigger thresholds and diverse discriminators.
- Readout: The end product of `sim_telarray` is the digitized pulse shape of the PMTs along with their calibration parameters such as pedestals, photo-electron conversion factors or flat-fielding coefficients. Triggered events are digitized and individual pulses are recorded over a wide time window (~ 100 ns) in order to compare different signal extraction approaches.

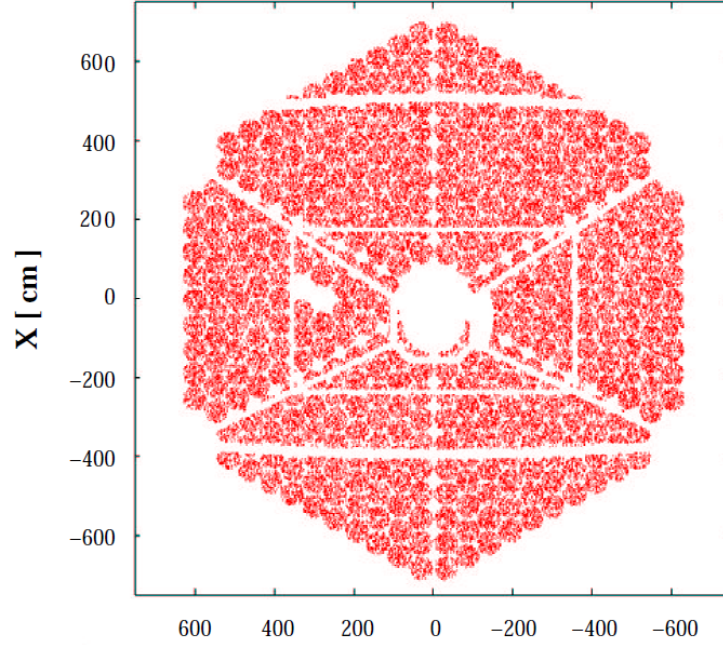


Figure 3.7: Arrival position of Cherenkov photons in the reflectors, reaching the camera plane. It shows how detailed the ray tracing is, displaying the masts and camera lid shadowing effect of the MST. Extracted from `sim_telarray` reference manual by K. Bernlöhr.

`sim_telarray` allows to evaluate telescopes response up to an incredible degree of precision. Evaluating the CTA generated CORSIKA output files, it simulates independently every telescope present in the super layout of telescopes introduced in Sec. 3.2.1 (different for each MC production) applying certain criteria in order to remove unnecessary data and improve efficiency, such as requiring stereo observation or a minimum amount of Cherenkov photons to perform the ray tracing.

The output is intended to be as similar as possible to the real IACTs *raw data* format. Once this data is generated, MC simulations are concluded, and the only missing step to estimate CTA performance is the MC analysis.

3.3 CTA Analysis tools

Starting at the pulse shape of the PMTs of each telescope camera, a complete analysis is required in order to properly evaluate CTA capabilities. The analysis performed to the simulated data is analogous to the one applied in current IACT experiments, briefly introduced in Sec. 3.3.1.

Taking into account the huge computational effort needed to perform the MC simulations, there is no possibility for cross-checking the entire telescope response. But the resulting data should be analyzed by as many independent methods as possible to ensure the validity of inferred conclusions. For this reason, the development of alternative analysis tools allows to reach independent conclusions by using different approaches, helping to accurately estimate CTA capabilities and also provide experience to improve the future CTA analysis software.

Several parallel approaches of the CTA analysis were successfully applied by different groups, generally using the software packages employed in the current IACT experiments. A brief introduction to the different methods used to analyze CTA MC simulations can be found in Sec. 3.3.2.

3.3.1 Data analysis overview

In this section, a general overview of the classical data analysis used in the IACT experiments is performed, introducing the main steps that need to be performed. More specific details on the analysis applied in this work can be found in Sec. 3.3.3.

The goal of data analysis is to effectively characterize the γ -ray events coming from an astronomical source to detect and measure its emission. To do so, the original particle direction and energy are estimated, suppressing as many background as possible. Once events are characterized, a subset of the total data is selected by applying certain cuts to attain the highest γ -ray efficiency as possible. The main cuts applied take into account differences between γ -ray and hadronic generated EASs and the arrival directions to be consistent with the position of the observed source. These cuts are then applied to the events falling within the *ON* region, corresponding to the collected data from the expected source position, and the

ones within the *OFF* region, conformed by the sky areas where no γ -ray emission is expected. Labeling the number of surviving events of these regions as N_{on} and N_{off} respectively, the excess events are defined as follows:

$$N_{exc} = N_{on} - \alpha N_{off}, \quad (3.1)$$

where α is the time ratio between *on* and *off* observations, used as a normalization factor. Note N_{on} is composed both by γ -rays and background, and αN_{off} (also referred as background events N_{bkg}) is the amount of background events expected to be within N_{on} . With these values, a statistical test is performed to calculate the significance level of detection S of the observed source (how likely observed *on/off* regions are consistent with no γ -ray emission). In γ -ray astronomy, significance is usually calculated using the expression (Eq. 17 from [136]):

$$S = \sqrt{2} \left\{ N_{on} \ln \left[\frac{1 + \alpha}{\alpha} \left(\frac{N_{on}}{N_{on} + N_{off}} \right) \right] + N_{off} \ln \left[(1 + \alpha) \left(\frac{N_{off}}{N_{on} + N_{off}} \right) \right] \right\}^{\frac{1}{2}}. \quad (3.2)$$

Note that as a convention, the detection of a source can be stated if its significance level equals or surpasses 5 standard deviations ($S > 5 \sigma$). If several trials are used to test a detection (for example, using several data cuts), this value needs to be increased in order to avoid false detections, taking into account the trials factor.

There are differences between the analysis of different IACT experiments, but there are some common steps they follow:

- **Calibration:** It comprises the methods used to convert individual PMT responses into a calibrated integrated charge in photo-electron counts and their corresponding arrival times for every pixel of the camera.
- **Image cleaning:** Recorded showers images by IACTs have two main components: the Cherenkov photons generated in the EAS and the NSBL. Cleaning algorithms are applied to suppress this second component, and ideally keeping all measured Cherenkov photons.

3. SENSITIVITY STUDIES FOR THE CTA

- Shower parameterization: Once camera images are cleaned, their resulting shower components are parameterized using the *Hillas parameters* [127].
- Data selection: With shower images parameterized, its possible to compare them to the averaged values expected in each observation as data is heavily affected by bad weather and sky conditions or telescope hardware malfunction. Only high quality data is usually selected to be used in the following steps of the analysis.
- Event characterization: Global event information is extracted from individual shower images inferring the most important parameters from the primary particle: direction, estimated energy and its likelihood to represent a γ -ray induced shower.
- Significance calculation: After the event characterization, a significance level of detection is calculated from the number of excess events in the *on* and *off* regions using Eq. 3.2.
- Higher level products: If the γ -ray source is positively detected ($S \geq 5 \sigma$), higher level products are generated such as the flux spectrum, SEDs, light curves or sky maps. In the case there is no positive detection, upper limits to the source flux can be computed.

During this process, there are common steps used in order to reduce the overwhelming population of background events that dominate collected data. First, the geometrical information of the observed images of each event is used to estimate how likely it is to be a γ -ray induced EAS, selecting only γ -like events. To further suppress background from collected data, the reconstructed arrival direction is used to only select events agreeing with the location of the expected γ -ray emitting source. To do so, an additional data cut is applied taking into account the squared angular distance between the location of the reconstructed events origin and the astronomical γ -ray source position in the sky (θ^2).

3.3.2 Alternative CTA MC analysis methods

As previously introduced, parallel independent analysis methods applied to the CTA MC data are crucial to ensure the validity of inferred conclusions, and their comparison allows to evaluate the effect of different event characterization approaches to improve future CTA analysis software.

All the analysis methods successfully applied by the different groups belonging to the CTA Consortium, using software validated within the current IACT experiments, are the following:

- **Baseline analysis:** This analysis uses the most simplified approach of the IACT analysis. After signal extraction, a 2 level cleaning [137] is performed and images are parametrized with the *Hillas parameters*. Stereo reconstruction is determined by a weighted average of direction intersection pairs from valid telescopes, weighting in terms of the projected stereo angle, *size* and *width* over *length*. Energy is estimated for individual images using LUTs, trained with a subset of simulated γ -rays, as a function of the *core distance* and *size*. The energy of the primary particle is estimated with a *size* weighted average of these values. Background rejection is performed using selection cuts, both using image shape and shower direction parameters. Mainly developed by K. Bernlöhner from the Max-Planck-Institut für Kernphysik (MPIK).
- ***evndisplay* analysis:** *evndisplay* is a software package used by the VERITAS collaboration. It performs the trace integration and a 2-level cleaning before calculating the second moment analysis. To decrease the effect on *leakage* (partially contained images in the camera) a likelihood fitting algorithm is applied, improving event reconstruction at larger *core distances* and therefore the sensitivity. Stereo reconstruction is performed using the algorithm number 1 from [138] and primary particle direction is calculated from the mean intersection points between all possible reconstructed image pairs. Energy is reconstructed by using a LUT as a function of the impact parameter, size, camera offset, level of NSB, zenith angle and telescope type. The energy is then calculated from the average of the estimates of each individual telescopes observing the shower. An analog approach is used to

3. SENSITIVITY STUDIES FOR THE CTA

calculate a mean-scaled width and length of the shower, parameters used to improve background rejection. The shower *maximum height* [139] is calculated from the weighted average values calculated from every image pairs. Background rejection is performed using MVA methods as a function of the following variables: shower direction, mean scaled width, the second largest image size, the maximum height, and the χ^2 value of the energy and maximum height calculation. The CTA *evndisplay* tools are mainly developed by G. Maier from the Deutsches Elektronen-Synchrotron (DESY).

- **SLAC analysis:** This analysis is mainly designed to estimate SC-MST performance [140]. It is the only analysis that uses Cherenkov photons directly (CORSIKA output), simulating a simplified telescope response and light collection efficiency with no ray tracing. Image cleaning is performed using the *aperture cleaning*, applying cleaning thresholds on the integrated signal in circular regions around pixels, needed for efficient cleaning of highly pixelated images. Shower reconstruction is performed in two stages: First, Hillas parameterization is carried out. Then, calculated values are used for a likelihood template analysis for the reconstruction of the primary particle direction and energy. Background rejection is performed using MVA methods as a function of the mean scaled image parameters, *maximum height* and the image goodness of fit (GoF) which measures the agreement between observed images and simulated ones. These tools were mainly developed by M. Wood from the Stanford Linear Accelerator Center (SLAC).
- **IFAE analysis:** Based on the MARS software package used by the MAGIC collaboration. This analysis is the one employed to analyze the CTA MC data in this work, and will be explained with a deeper scope in Sec. 3.3.3. The trace integration is performed in a 2 step process, using the arrival time to decrease the NSB level. Two level absolute cleaning and Hillas parameterization is performed. Direction reconstruction is calculated minimizing the shower direction distance, weighted with the reconstructed direction deviation from a look-up-table as a function of *size*, *width* over *length* and telescope type. The primary particle energy is inferred from a weighted mean of shower estimated energies, calculated using MVA methods. Gamma-hadron

separation is also calculated from a weighted mean using MVA, using both shower and stereoscopic parameter. These tools were mainly developed by A. Moralejo from the Institut de Fisica d'Altes Energies (IFAE).

For the proper comparison of the different analysis tools and the usability of their results, the *evndisplay*, the SLAC and the IFAE analysis products are stored in the standard performance root file defined by the CTA Consortium.

3.3.3 CTA MARS based analysis

In this work, the CTA MC generated data is analyzed using MAGIC Analysis and Reconstruction Software (MARS) [141, 142], the software package used by the MAGIC collaboration. This software is written in C++ and uses the ROOT data analysis framework [143]. The MARS package is developed to analyze the MAGIC data products, and needed to be adapted to handle CTA MC format. A new set of analysis tools, detailed in the following sections, were developed to provide good performance estimators for the CTA.

The different steps within the CTA MARS based analysis are the following:

- Trace integration: Pixel charge and the time of arrival are extracted from `sim_telarray` output for each triggered event. Converts the output to the MARS format (ROOT format).
- Image cleaning and parameterization: Image cleaning attempts to remove NSBL and retain Cherenkov photons using a two level absolute cleaning. Subsequently, parameters describing the shower image geometry are calculated.
- Stereo reconstruction: Using the information gathered by all triggered telescopes observing an event, stereo parameterization is performed along with the energy and direction estimation, and how likely the event is a γ -ray or background.
- Performance estimation: The MC events are weighted and optimal cuts determined in order to calculate the resulting performance, expressed by the Instrument Respond Functions (IRFs).

3. SENSITIVITY STUDIES FOR THE CTA

3.3.3.1 Trace integration

The first step of the analysis is to convert CTA MC data into the MARS format. The Convert Hessio Into Mars inPut (CHIMP) package transforms `sim.telarray` output into calibrated data stored in ROOT files using HESSio libraries [135]. It performs a 2-pass signal extraction method for each pixel of each triggered event, maximizing the acquisition of Cherenkov photons and decreasing the integrated NSBL.

The simulated Flash analog to digital converter (FADC) counts from `sim.telarray` output are stored in a different format for each telescope type. As low energy events show shorter peaks, the LST is designed to integrate narrow time lapses and only a 30 ns lapse is stored (over 30 slices of 1 ns). This time trace increases for telescope types prepared to measure higher energy events, reaching values as high as 160 ns (over 40 slices of 4 ns) for 4m DC-SSTs.

Signal is extracted in a 2-pass process, following these steps for each triggered telescope camera:

- Sliding window extractor: For each individual pixel of the camera, the extractor searches for the maximum sum of consecutive samples using a fixed time window through the whole range of the FADC slices. Signal is calculated subtracting the mean pedestal value from the highest integral, while the time of arrival is determined through a weighted average of the window sample. The width of the sliding window is optimized for every telescope type, and ranges between 4 ns to ~ 15 ns.
- Cleaning and time fit: In order to infer the time of arrival of the Cherenkov photons along the camera pixels, a time fit is performed to the cleaned image. The standard 2 level cleaning is applied (detailed in Sec. 3.3.3.2) removing most of the NSBL dominated pixels, and fit the arrival time of the surviving ones as a function of the distance to the Center of Gravity (CoG) along the shower major axis. Only pixels belonging to the main *island* are used in the fit.
- Second pass extraction: If the fit shows no correlation with the extracted arrival times, signals recorded in the first pass are stored. If the fit is accept-

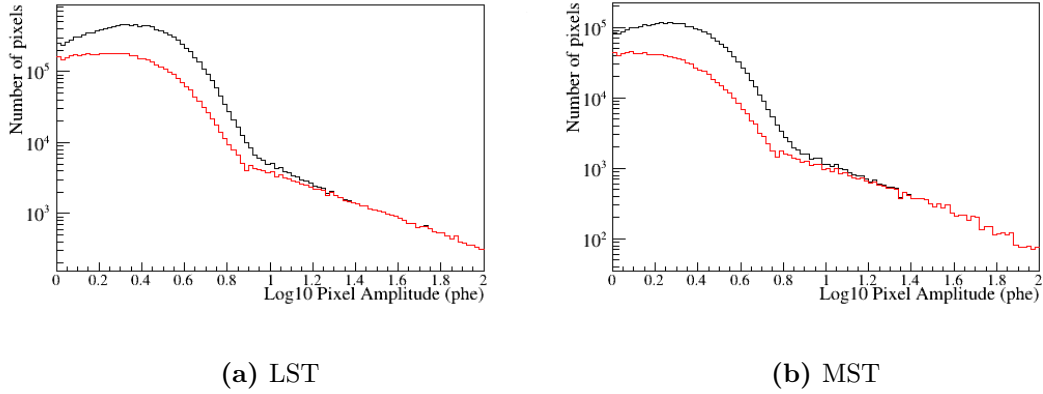


Figure 3.8: Histograms of the logarithm of the pixel amplitude (extracted charge from individual pixels) of two different CTA telescope types: the LST and the MST. The *Black lines* show the extracted signal in the first pass, and the *red lines* show it after the second pass, using information from the Cherenkov photons time of arrival.

able, a second pass is performed taking this information into account. Pixels not used in the time fit (not surviving the image cleaning) are re-extracted placing the integration window at the corresponding time of the fit.

Note that the time fit may not be applicable to every event. Showers are required to be intense and compact enough to survive image cleaning and 3 or more pixels need to belong to the main island.

Using this method, only NSB photons randomly entering the integration window at the corresponding time are stored, substantially decreasing the signal to noise ratio. As shown in Fig. 3.8, NSBL is heavily subtracted by the 2-pass extraction method (corresponding to low amplitude pixels) while not having a significant effect on Cherenkov photons dominated pixels (high amplitude pixels). It must also be noted that the time fit information is recorded and used in later steps of the analysis as the time Root Mean Squared (RMS) with respect to the fit turns out to be an useful parameter to gauge image quality.

3.3.3.2 Image cleaning and parameterization

After the 2-pass extraction process, telescope images are characterized by the arrival time and the pixel charge, proportional to the number of photo-electrons

3. SENSITIVITY STUDIES FOR THE CTA

measured by the PMT. In order to accurately calculate the *Hillas parameters*, the NSBL noise in the camera needs to be suppressed further, using image cleaning algorithms.

Within the different image cleaning algorithms, this software uses a 2 level absolute cleaning. In this method, surviving pixels belong to two different populations: *core* pixels and *boundary* pixels. Neighboring pixels with signals above the first cleaning level q_{core} (in *phe* units) are tagged as *core* pixels, forming the *core* of the image. Adjacent pixels to the first population with signals between the first and the second cleaning levels ($q_{core} > q_i > q_{bound}$) are considered boundary pixels. The rest of the camera pixels are removed. The cleaning parameters used for the different CTA telescopes can be found in Table 3.1.

As shown in Fig. 3.9, Hillas parameters are highly affected by the effect of NSBL in the shower image. Increasing cleaning levels show how spurious photons are subtracted, improving parametrization and hence direction reconstruction.

Table 3.1: Trace integration and image cleaning parameters for the CTA telescopes.

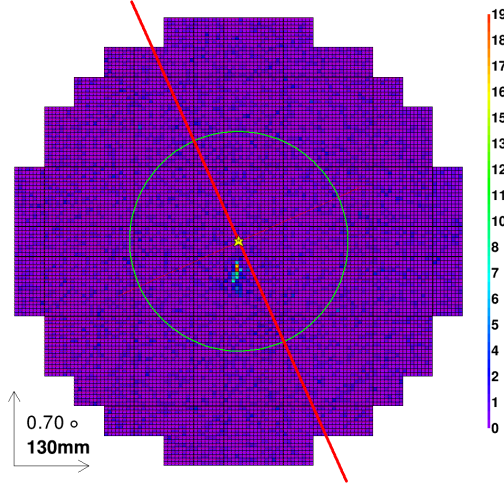
	LST	MST	SC-MST	DC-SST	Schwarzschild-Couder (SC)-SST
Pulse width [ns]	4	8	10	12	12
q_{core} [phe]	4	6	2.5	6	3
q_{bound} [phe]	2	3	1.25	3	1.5

Width of the sliding window in the 2-pass trace integration process and *core* and *boundary* image cleaning thresholds for every telescope type present in the second large-scale CTA MC production.

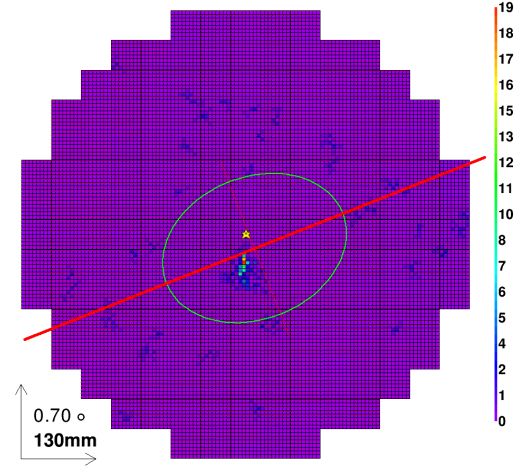
Note that this cleaning method has been used by the MAGIC collaboration and it was designed to work with hexagonal pixels (6 neighbours per pixel) and cameras of ~ 1500 pixels. For densely pixelated cameras of square geometry (8 neighbours per pixel), like the one built-in the SC-MST, this cleaning method is not ideal, and other algorithms like sum cleaning [144] or aperture cleaning [140] show better performance.

Once the image cleaning is performed, individual shower images are parameterized using quasi classical *Hillas parameters*. The most important monoscopic (source independent) parameters in the analysis are the following:

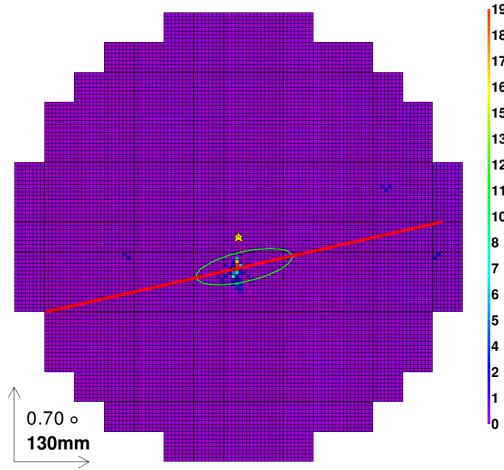
- *Size*: Sum of the total charge of the cleaned shower image. It is crucial in



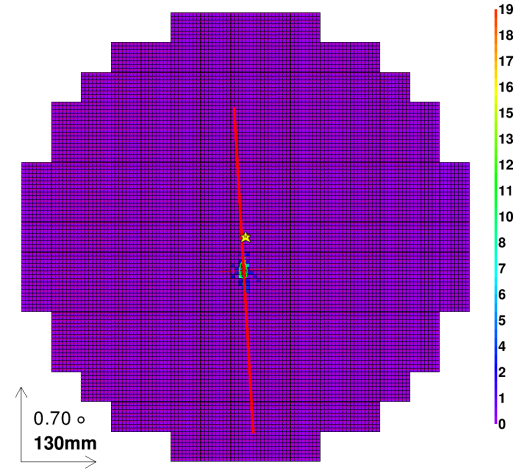
(a) No image cleaning.



(b) $q_{core} = 1$; $q_{bound} = 0.5$



(c) $q_{core} = 1.5$; $q_{bound} = 0.75$



(d) $q_{core} = 2$; $q_{bound} = 1$

Figure 3.9: Effect on the direction reconstruction of different image cleaning levels with the Hillas ellipse parameterization superimposed. The source position is marked with a star symbol in the center of the camera. Images correspond to a MC generated γ -ray of $E = 160\text{GeV}$ and $impact = 75\text{ m}$ triggered by a SC-MST.

3. SENSITIVITY STUDIES FOR THE CTA

every step of the analysis, specially energy estimation.

- *Width*: Distance (in mm) of the minor axis of the image. Together with the *length* parameter, they measure the longitudinal development of the EAS, and are crucial for γ -hadron separation.
- *Length*: Distance (in mm) of the major axis of the image.
- *Conc*: Ratio between the sum of the two highest pixels of the image and the total *size*. It measures how compact is the emission of the Cherenkov photons of the shower, usually larger for γ -ray events.
- *Number of islands*: Number of isolated pixel groups surviving image cleaning. Hadronic showers are generally more disperse, and show higher number of islands.
- *meanX(Y)*: Position in mm of the X (and Y) coordinates of the image Center of Gravity (CoG) (the center of the ellipse). Showers centered near the camera edges are not used, in order to remove showers with high *leakage*, which accounts for the fraction of the shower image that falls out of the camera.
- *RMSTimeFit*: RMS of the time distribution with respect to the shower major axis. It is used to suppress events with poor likelihood in the time fit performed during the trace integration, improving data quality.

In Fig. 3.10 a schematic view of the monoscopic parameterization is pictured, showing the fitted ellipse in green centered in the *mean*, with the *width* and *length* distances highlighted. All γ -ray showers major axis (red line) should be aligned with the position of the γ -ray source projected into the camera (star symbol).

3.3.3.3 Stereo reconstruction

The next step in the analysis is to characterize EASs event-wise, inferring information of the primary particle from the images gathered by the triggered telescopes. The main characteristics to determine are the primary impinging direction, its energy and to discern if its nature is of γ -ray origin.

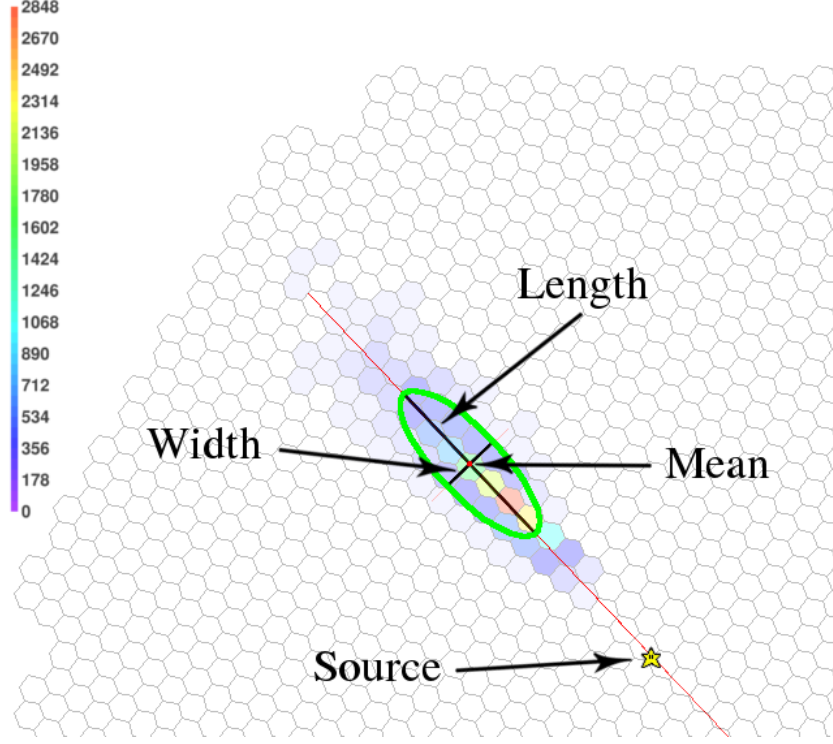


Figure 3.10: Schematics of the monoscopic parameterization using the first order Hillas parameters. Image corresponds to a simulated γ -ray of $E \approx 3 \text{ TeV}$ seen by a LST.

Direction reconstruction: The primary particle direction is calculated from individual telescope images, determining the source position in the camera by minimizing $\sum_n w_i d_i^2$, where d_i is the distances between each shower direction and the source location and w_i the weight of each telescope image. This weight accounts for the quality of the reconstruction considering the shower geometry and the telescope type. It is extracted from a previously generated LUT from an independent set of simulated γ -rays, evaluating the reconstructed deviation (mismatch between reconstructed and real source position) as a function of *size* and *width* over *length*. Then, the *core position* and the shower *maximum height* are calculated using similar minimizations. Note only triggered telescopes with clean images (reduced *Number of islands*) and *size* larger than a certain threshold (50 phe) are used in the minimization.

The following stereoscopic parameters are then calculated event-wise:

3. SENSITIVITY STUDIES FOR THE CTA

- *Core position*: Position in the ground intersecting the projection of the primary particle direction, corresponding to the center of the Cherenkov light pool.
- *Maximum height*: Estimated height (in cm) at which the maximum development of the shower was located. This parameter is used for the energy reconstruction and also by the background suppression algorithms.
- *Reconstructed source position*: Point projected in the sky (in degrees) which represents the estimated position of the source emitting the event (considering the event as a γ -ray).
- *Multiplicity*: Number of telescopes used in the direction reconstruction.
- *MinResultDir*: Minimum value (in square degrees) obtained for the minimized function ($\sum_n w_i d_i^2$) during the direction reconstruction.

And the following monoscopic (source dependent) parameters are calculated image-wise:

- θ : Distance (in degrees) between the *reconstructed source position* and the position of the γ -ray source. This parameters squared, together with the *hadroness*, is the main background suppression cut generally applied in the IACT technique.
- *Impact parameter*: Distance projected in the ground (in cm) between the core position and the telescope observing the event.
- *Dist*: Angular distance (in degrees) between the center of gravity of the image (*mean*) and the reconstructed source position in the camera.

Energy estimation: Two different methods are implemented to estimate the energy of the primary particle, always considering them of γ -ray nature. The first method is analog to the one implemented for the direction reconstruction, and uses a Look-Up Table (LUT) (one per telescope type) evaluating reconstructed showers *energy* over *size* as a function of the *impact* parameter and the stereo

reconstructed *maximum height*. The second method used is a Random Forest (RF) (also one per telescope type) trained to evaluate the energy as a function of a list of parameters: The stereo reconstructed *maximum height* and the monoscopic *size*, *impact* parameter, *width*, *length*, *concentration ratio*, $\frac{width*length}{size}$, $dist^2$ and the angle between the positive x-axis and the line projected by the source position and the CoG of the shower. Both methods need a previous step computing the LUT or the RF using an independent set of generated γ -rays, and they are both used to estimate the energy of individual images independently. The primary particle energy estimation is calculated as the average of these values weighted with $w_i = \frac{1}{\Delta E^2}$. The performance of these two methods is compared in Sec. 5.2.1.

Gamma-Hadron separation: Background rejection is performed by applying a RF algorithm for each telescope type, trained to discern between showers of electromagnetic or hadronic origin. This algorithm is applied to each triggered telescope image assigning a real number, defined between $[0, 1]$, called *hadroness* (h). This variable indicates the likelihood of a shower to have hadronic origin corresponding $h \approx 1$ to hadron-like showers and $h \approx 0$ to γ -like ones. The RF is trained using both simulated γ -rays and hadrons, evaluating *hadroness* as a function of the stereo estimated *energy* and *maximum height*, and individual telescope images *energy*, *impact* parameter, *size*, *width*, *length* and *concentration ratio*. The global hadroness value is calculated as the average of the calculated single values weighted with $w_i = size^{0.54}$ (expression obtained empirically).

Taking into account the methods just described, 4 different datasets are required. An initial set of γ -rays is needed to train the LUT involved in the direction reconstruction. In principle, a training sample of $\sim 2 \times 10^5$ stereo triggered shower images would be enough for accurate direction reconstruction weighting. A second set of stereo reconstructed γ -rays is used for the LUT (or RF) involved in the energy reconstruction. This sample also needs $> 2 \times 10^5$ telescope images to create a proper training. The third dataset corresponds to the γ -rays and hadrons needed for the RF training, responsible of the background suppression. The number of telescope images needed for a decent cosmic-ray rejection is $\sim 1 \times 10^5$ for each telescope type, both for γ -rays and hadrons ($\sim 2 \times 10^5$ in total). Once all

3. SENSITIVITY STUDIES FOR THE CTA

the different methods needed for the event characterization are computed, the last dataset (composed by the remaining MC simulated γ -rays, hadrons and electrons available) is used to estimate the CTA performance, explained in the next section.

Note that in the last CTA MC production, stereoscopically triggered γ -ray events account roughly for a $\sim 0.3\%$ of generated showers, generally observed by ~ 4 telescopes (although this value increases significantly with the energy). In the case of generated hadrons, this number significantly decreases, accounting triggered events for less than the $\sim 0.02\%$ of generated hadronic showers.

Instead of using all triggered telescopes to characterize an event, a set of quality cuts are applied to get rid of poorly reconstructed showers, significantly improving the characterization. Showers are required to exceed a minimum *size*, *main island size over size* ratio and *width over length* values, removing faint and scattered images. The dispersion from the 2-pass trace extraction time fit and the number of islands are also used to reject noise dominated events.

3.3.3.4 Performance estimation

With simulated events characterized, the performance of the observatory needs to be estimated by optimizing the different parameters involved, maximizing the differential sensitivity.

Data products (γ -rays, protons and electrons) are weighted in order to resemble realistic fluxes: The γ -ray population is weighted using the Crab Nebula spectrum (by HEGRA [145]), protons using the cosmic-rays spectrum measured by BESS [146] and electrons using a parameterization of the *Fermi* and H.E.S.S. spectrum data.

The differential sensitivity is defined as the minimum flux from a steady γ -ray source observable by the observatory in a certain period of observation time. Classically, IACT experiments use the differential sensitivity of an steady point-like γ -ray source in 50 hours of observation as the standard parameter to assess their performance.

In order to optimize sensitivity for a given observation time, optimal cuts for 3 different parameters are calculated between 20 GeV up to 200 TeV: *hadronness*, θ^2 and *multiplicity*. These cuts need to find the equilibrium between different cut

efficiencies (ratio of surviving γ -rays after each cut) and the number of remaining γ -rays. Very restrictive cuts would increase the signal to noise ratio, but reduce the needed effective area to significantly observe a γ -ray source in the selected observation time. The *multiplicity* cut applied corresponds to the minimum number of telescopes used in the stereoscopic direction reconstruction of each event, ranging between 2 to 9. As an example, a *multiplicity* cut of 3 would reject all events using only 2 telescopes in the stereo reconstruction, and surviving events with 3 or more telescopes. Note that these samples are not independent, as events with higher *multiplicity* values are also present in lower *multiplicity* cut populations.

The optimization is carried out as follows: *hadroness* and θ^2 cuts are determined by calculating the corresponding sensitivity of several cuts with fixed increasing efficiencies, going from the highest (less restrictive) to the lowest values considered (50% of surviving γ -rays). This process is performed for each energy bin and each *multiplicity* cut considered. Finally, the best sensitivity of each energy and *multiplicity* out of each cut efficiency is selected, determining this way the cuts maximizing sensitivity. Note the minimum observable flux is calculated using standard detectability conditions, with a minimum statistical significance of 5σ , at least 10 γ -ray events and the excess is required to be greater than a fraction 0.05 of the residual background (considering 5% of systematics). Fig. 3.11 pictures the minimum observable flux for different *multiplicity* values, showing high multiplicities are more efficient in the core energy range of the CTA (improved stereo reconstruction), while smaller values improve the low (due to the reduced number of LSTs) and very high energies (to maximize the effective area).

Different observation modes affecting performance, such as different pointing directions need to be evaluated. The methods described in this section allow several sub-samples to be analyzed independently and properly re-weighted to obtain the averaged performance over several data samples. This is generally used to calculate the CTA performance evaluating both North and South pointing directions (significantly different due to the different Geo-magnetic field effect over the EASs). Note this sub-samples need to be analyzed with independent training (including different LUTs and RFs) and testing samples.

Besides the performance on a point-like steady γ -ray source located in the center of the camera, its also important to estimate CTA off-axis capabilities in order

3. SENSITIVITY STUDIES FOR THE CTA

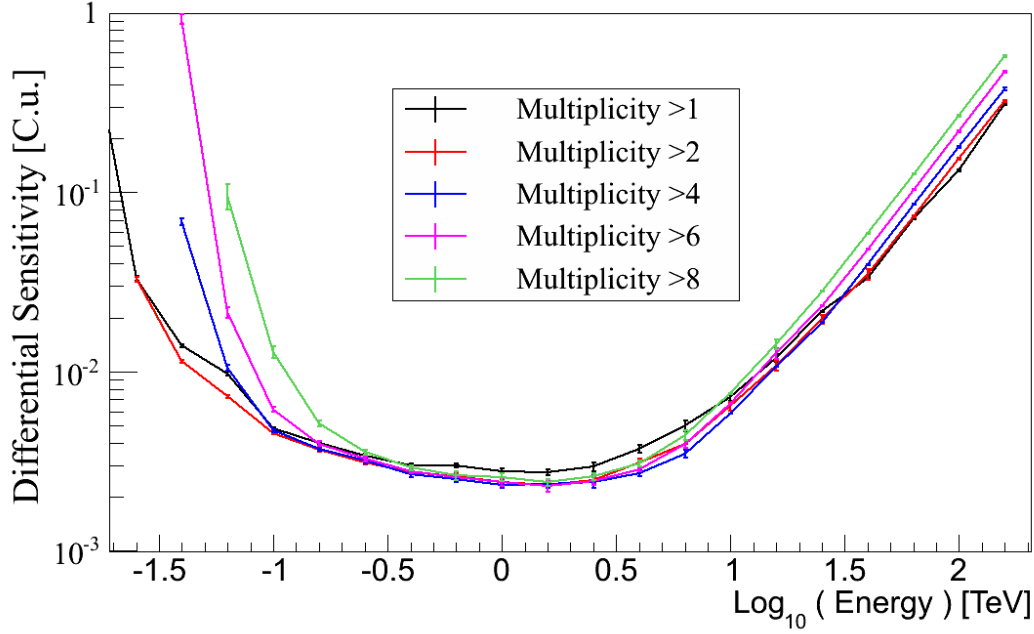


Figure 3.11: Differential sensitivity (in Crab units) in 50 hours for the “2Q” candidate layout simulated at Namibia for 5 different *multiplicity* cuts.

to take advantage of its wide FoV. These methods allow the analysis of CORSIKA generated diffuse γ -rays, evaluating CTA sensitivity of a source located at different distances from the center of the camera. Both diffuse γ -rays and background events are separated taking into account the off-axis angle, following an equivalent event characterization as the on-axis analysis, with the peculiarity of calculating diffuse γ -rays θ^2 values with respect to the actual true direction of the primary particle. As the generated diffuse γ -ray statistics are very limited, cut efficiencies are extrapolated from the on-axis values, while *hadroness* and *multiplicity* cuts are looped to maximize differential sensitivity for each different off-axis angle. Note this analysis still needs to be improved: it suffers significantly from the lack of statistics caused by the distribution of the simulated events among the off-axis binning, specially at high off-axis angles.

Results are stored in the standard performance root file, defined by the CTA collaboration, containing the main IRFs of the observatory: Differential sensitivity, angular and energy resolutions, effective area, background rates and the migration

matrix. These files are then used by other software tools to produce estimations of higher level products such as measured spectra, light curves, upper limits or skymaps. See Sec. 4.1 for an overview of these tools.

3.4 CTA Monte Carlo productions

As briefly introduced in Sec. 3.2.1, due to the great computational effort needed to obtain coherent performance estimations for such large scale layouts with such a huge energy coverage, few large-scale MC productions have been performed. Each of these productions define a different large telescope layout and update their configuration to parameters as realistic as possible.

During the early stages of the CTA collaboration, preliminary simulations were performed of some telescope arrays comprising different size and layout shapes. Some of these first attempts were: the so called *benchmark array* with 13 LST-like telescopes simulated at 3 different altitudes (2000, 3500 and 5000 m), a large-scale layout of 97 H.E.S.S.-like telescopes and several layouts with 7-m of diameter SST-like telescopes. These simulations were mainly driven by K. Bernlöhr, and helped to define the first large-scale production.

Note each large-scale MC production requires the simulation of an enormous number of showers, initiated by both γ -rays (point-like and diffuse) and cosmic rays (mainly protons and electrons).

3.4.1 Prod-1

The Prod-1 was the first CTA large-scale MC production, simulating a super-layout of 275 telescopes: 10 LSTs of 23m diameter in the central region, 125 DC-MSTs of 12m diameter (or simply MST from now on), 96 SSTs of 7m diameter, 40 MSTs of 12m diameter with an increased FoV and more pixelated cameras and 4 MSTs of 10m diameter with wider FoV. The distribution of simulated telescopes is shown in Fig. 3.12 and detailed telescope parameters are listed in Table 3.2. This production also allowed the development, testing and comparison of most of the analysis tools currently used for the CTA MC data. Prod-1 was computed on the Grid at MPIK and INTEGRAL Science Data Centre (ISDC).

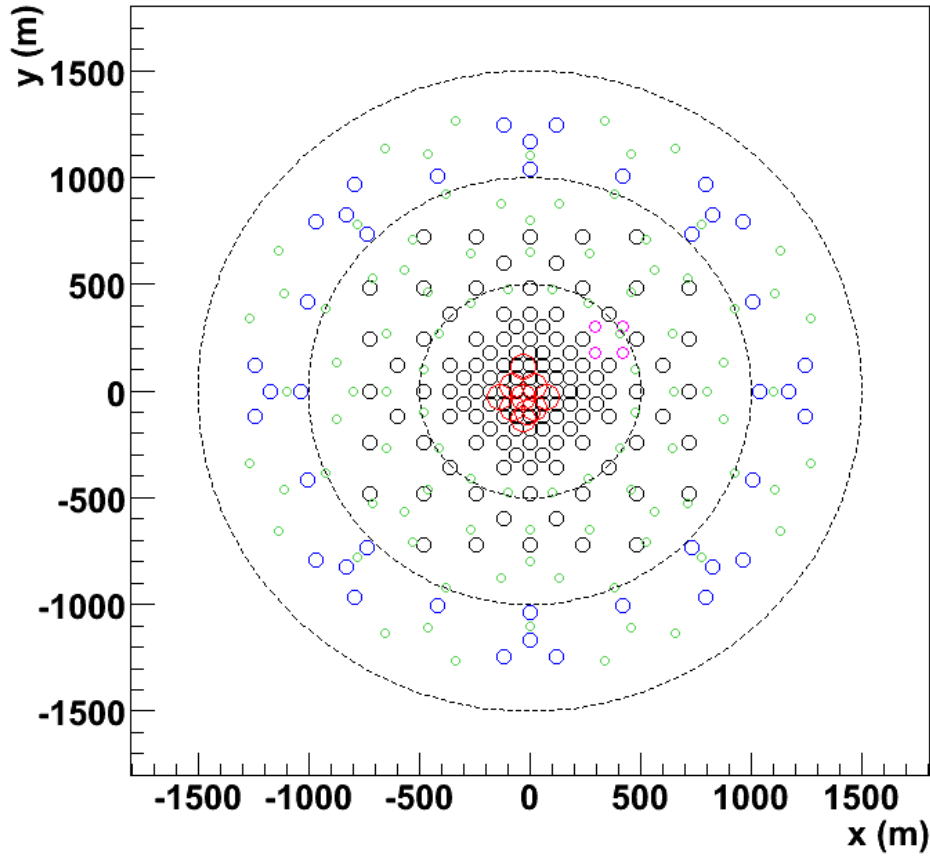


Figure 3.12: Complete layout of simulated telescope positions of the first large-scale CTA MC production (Prod-1). Telescope types (and diameter) correspond to the following colors: *Red*: 23 m LSTs. *Black*: 12 m MSTs. *Pink*: 10 m “Test” MSTs. *Green*: 6.7 m SSTs. *Blue*: 12 m “WF” MSTs with 10° of FoV.

Table 3.2: Parameters of Prod-1 simulated telescope.

	Diam.	Mirror A.	Focal l.	Pixels	Pix. Diam.	FoV	NSB
	[<i>m</i>]	[<i>m</i> ²]	[<i>m</i>]		[<i>cm</i>]/[<i>deg</i>]	[<i>deg</i>]	[<i>MHz</i>]
LST	23	412	31.2	2841	4.9 / 0.09	5	122/180/220
MST	12	100	15.6	1765	4.9 / 0.18	8	120/178
SST	6.7	37	11.2	1417	4.9 / 0.25	10	85
WF MST	12	100	15.6	2737	4.9 / 0.18	10	120
Test MST	10	73	10.0	931	5.2 / 0.30	10	237

Using such telescope distribution, a wide variety of conditions were simulated using a constant geomagnetic field strength (averaged from candidate sites): two different altitudes (2000m and 3700m), different source positions at zero, medium and large zenith angles (0°, 20° and 50°) and partial moon light.

The goal of Prod-1 was to give a first accurate estimation of many CTA candidate configurations, evaluating very different telescope distributions of approximately equal cost. Therefore a set of ~ 10 candidate layouts were defined and used to test different physics cases by the Consortium and decide which layouts would have a stronger scientific impact. Some of these studies are detailed in chapter 4. Table 3.3 shows the different number and type of telescopes within each of the proposed layouts, and Fig. 3.13 shows several examples. The subsets B and D show extreme layouts with performance focused on the low and high energy ranges respectively, while layouts E and I are more balanced options.

Table 3.3: Prod-1 candidate layouts

	NA	NB	A	B	C	D	E	F	G	H	I	J	K
LST	4	3	3	5	-	-	4	6	6	-	3	3	5
MST	17	17	41	37	29	41	23	29	9	25	18	30	-
SST	-	8	-	-	26*	16*	32	-	16	48	56	16*	72

Telescope composition of Prod-1 proposed array candidates for the CTA North (NA, NB) and CTA South (layouts A to K). SSTs marked with a * correspond to WF-MSTs.

The main conclusions of Prod-1 were the following:

3. SENSITIVITY STUDIES FOR THE CTA

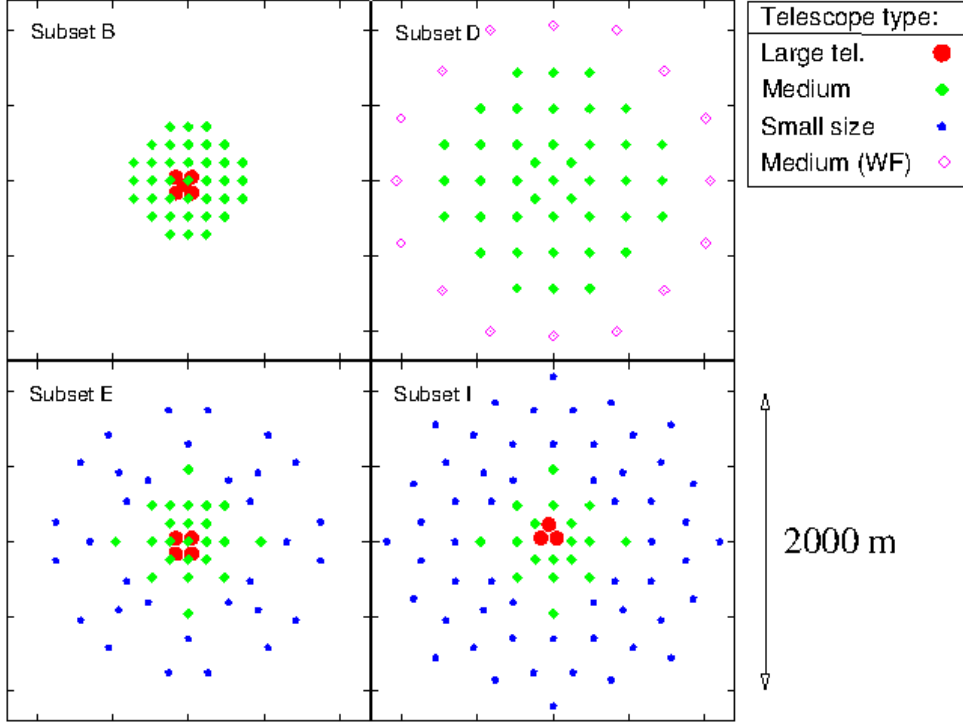


Figure 3.13: Telescope distribution of 4 proposed candidate arrays from the Prod-1. Layouts B and D represent different approaches to focus on the low and high end of the energy range respectively. Layouts E and I are the most balanced examples, with an uniform sensitivity from 20 GeV up to 200 TeV.

- Several of the proposed layouts could satisfy CTA scientific expectations for most physics cases.
- Balanced proposed array candidates like layout E and I resulted favoured, taking into account their potential scientific outcome.
- The improvement at the lowest energies ($E < 100 \text{ GeV}$) of high altitude building sites does not compensate the effective area loss above those energies, favouring sites at moderate altitudes.

A parallel production was performed named *Hybrid-1*, intended to simulate both DC-MSTs and SC-MSTs telescopes in a mixed layout assessing the true gain of using SC optics and mixed arrays performance. The same software for shower and telescope simulations was used as in Prod-1. It was used for the calibration

and development of analysis tools for the SC-MSTs, although final conclusions on sensitivity comparison were limited by the lack of proton statistics.

3.4.2 Prod-2

The second CTA large-scale MC production (Prod-2) intended to go one step further, with a clear main objective: to assess the impact of the construction site properties on the observatory sensitivity and the search for a feasible layout candidate, taking into account updated cost estimates and new proposed telescope types. With this in mind, telescope responses, site properties and the layout composition were configured with more realistic parameters:

Telescope response: Using all `sim_telarray` improvements and the updates from all the telescope designs (such as PMTs, mirror reflectivity or readout systems), full ray tracing was performed, recording long readout windows to allow custom trace integration by analyzers and 3 different trigger schemes: majority, analog-sum and digital-sum.

Layout composition: New large macro-layouts of telescopes are defined (one for each simulated Hemisphere) with new telescope positions and types. Also parallel productions with new super-layouts are simulated using the same CORSIKA showers, allowing to merge and analyze a wider variety of layouts.

Site properties: For both the Northern and Southern Hemisphere candidate sites, specific atmospheric density profiles, altitudes and geo-magnetic fields (direction and intensity) are used.

A wide range of telescope types and several super-layouts were simulated within the different parallel `sim_telarray` configurations within the Prod-2 in order to assess every telescope design performance with an unbiased comparison and to provide a wider range of candidate layouts. All Prod-2 configurations along their code names, listed in Table 3.4, are defined as follows:

3. SENSITIVITY STUDIES FOR THE CTA

- **STD:** Standard layout for the Southern Hemisphere comprised by 197 telescopes: 9 LSTs, 79 MSTs, 37 7m SSTs and 72 4m SC-SSTs. This configuration is used in every Southern Hemisphere simulated location, and sub-arrays of this super-layout (shown in Fig. 3.38, in the end of this chapter) are used to compare CTA-S sites performance. EASs are simulated over a circular area of 2500m and 3000m of radius for γ -rays and cosmic rays respectively.
- **NORTH:** Standard layout for the Northern Hemisphere comprised by 61 telescopes: 5 LSTs, 37 MSTs and 19 7m SSTs. This configuration is used in every Northern Hemisphere simulated location, and sub-arrays of this super-layout (shown in Fig. 3.39) are used to compare CTA-N sites performance. EASs are simulated over a circular area of 1600m and 2100m of radius for γ -rays and cosmic rays respectively.
- **SCMST:** Layout containing 111 SC-MSTs, impossible to add to the STD configuration due to memory restrictions. SC-MST positions are equal to the MSTs withing STD configuration, with an additional 32 positions (shown in Fig. 3.38). This configuration was simulated in several Southern Hemisphere sites: Leoncito, Leoncito++, Aar and Armazones_2K.
- **NSBx3:** Equivalent to STD, with an increased NSB by a factor 3. Telescope trigger thresholds are also adapted. Only simulated in “El Leoncito” site candidate.
- **4MSST, SCSST, ASTRI:** Extensions of 102 4m SSTs, generated to allow higher SST densities than with standard configuration, each corresponding to a different design of 4m SSTs: Davies Cotton small-sized telescopes (DC-SST), CHEC SC-SST and ASTRI SC-SST.

Table 3.4: Prod-2 Configurations

	LST	MST	SC-MST	7m SST	4m DC-SST	4m SC-SST CHEC	4m SC-SST ASTRI
STD	9	79	-	37	-	72	-
NORTH	5	37	-	19	-	-	-
SCMST	-	-	111	-	-	-	-
NSBx3	9	79	-	37	-	72	-
4MSST	-	-	-	-	102	-	-
SCSST	-	-	-	-	-	102	-
ASTRI	-	-	-	-	-	-	102

List of configurations produced within the Prod-2 along with the number and type of telescopes present in each of them.

In order to assess the impact of specific construction locations on CTA performance, a wide range of site candidates were simulated. Each location is simulated using different atmospheric density profiles, altitudes and geomagnetic fields. Note the different levels of NSBL affecting each region are not included in each site configuration, but is studied with an specific MC production (**NSBx3**). An alternative method to assess the NSBL effect in Tenerife site was done by the author of this work, detailed in Sec. 3.6.3. To understand the impact of these parameters to the EASs, and therefore to the CTA sensitivity (see [147]), we encourage the reader to leaf through Appendix A. The general effect of each of these parameters is the following:

- Atmospheric density profile: The overall effect is that different profiles impose the shower maximum of the EAS to be located at altitudes of different refraction index, affecting the amount of Cherenkov light emitted. As most of the Cherenkov light reaching ground close to the shower core is emitted at the shower maximum, different density profiles reduce the photon density in this region of the light pool. The atmospheric profile affect the overall performance of IACTs, specially at lower energies where photon density is crucial. These profiles are generated with MODTRAN [148].

3. SENSITIVITY STUDIES FOR THE CTA

- **Altitude:** There are two important effects concerning the altitude of construction, and could be considered double-edged. Placing IACTs closer to the shower maximum (higher altitudes) decrease the total area of the light pool (see Fig. A.4b from Appendix A) while increasing the Cherenkov photon density. This allows the collection of more Cherenkov photons by IACTs located near the shower core, but shrink the Cherenkov radius, significantly decreasing the light collected by IACTs farther away. Higher construction altitudes improve low energy performance but at the same time reduce the effective area at higher energies.
- **Geo-magnetic field effect:** The Geomagnetic Field (GF) bends the trajectory of charged particles generated within the EASs, broadening their lateral development, resulting in the loss of some Cherenkov photons decreasing collected light by IACTs close to the shower core. This effect also deforms the light pool on the ground, and shifts the γ -rays reconstructed direction affecting the angular resolution. This effect increases at higher zenith angles and is proportional to the orthogonal magnetic field intensity.

Within the Prod-2 the following candidate sites were simulated (throughout this work, their simulation code names will be used):

Northern Hemisphere:

- “US”: Simulation assessing both Meteor Crater (in Coconino County, Arizona) and Yavapai (Yavapai County, Arizona) North American sites.
- “SPM”: Corresponds to an observatory located in the Sierra de San Pedro Mártir National Park (Baja California), the Mexican proposed site.
- “Tenerife”: As the name stands, this simulation evaluates both Izaña observatory in Tenerife and can also represent the Roque de los Muchachos observatory in La Palma, the two Spanish proposed sites.

Southern Hemisphere:

-
- “SAC”: One of the proposed Argentinian site located at San Antonio de los Cobres.
 - “Armazones”: Corresponds to Cerro Armazones, the proposed Chilean site, simulated at two different altitudes: 2000m and 2500m.
 - “Aar”: Simulation that evaluates the two proposed Namibian sites: Aar farm and Göllschau (H.E.S.S. construction site). This site was also simulated at 500m to assess a lower altitude performance.
 - “Leoncito”: Corresponds to The Leoncito Astronomical Complex Argentinian site in the San Juan Province. The first proposed location was at an altitude of 2662m, and then modified to an alternative position at 1650m.

The simulated candidate sites along their location in geographic coordinates, altitude, and geomagnetic field intensity is shown in Table 3.5.

Table 3.5: Prod-2 simulated sites

Site	Location		Altitude [m]	B_h [μT]	B_z [μT]
	lon.	lat.			
	[deg]	[deg]			
US	-115.5	31	1655	23.5	42.9
SPM	-115.5	31	2434	25.2	38.3
Tenerife	-16.5	28.3	2290	30.8	23.2
SAC	-66.2	-24	3600	20.9	-8.9
Armazones	-70.2	-24.4	2500	21.4	-8.9
Armazones++	-70.2	-24.6	2000	21.4	-8.9
Aar	16.5	-26.7	1640	10.9	-24.9
Aar500	16.5	-26.7	500	10.9	-24.9
Leoncito	-63.3	-31.8	2662	19.9	-12.6
Leoncito++	-63.3	-31.8	1650	19.9	-12.6

List of simulated sites within the Prod-2 along with their geographic coordinates, altitude and the local GF intensity.

3. SENSITIVITY STUDIES FOR THE CTA

All Prod-2 simulations were performed at the MPIK by K. Bernlöhr and using the Grid with CTA resources, using 17 EGI sites spread through 6 countries, by L. Arrabito and J. Bregeon.

3.4.2.1 Layout candidates

Performing a similar exercise as in the Prod-1, and taking into account the produced super-layouts configurations, a wide variety of sub-samples may be selected as array candidates of approximately equal cost. Certain restrictions are imposed by the different configurations simulated in each site: The Northern Hemisphere comparison needs to be performed with a candidate array corresponding to a sub-sample of the **NORTH** configuration (with significantly less telescopes). Neither **SCMST** or any 4m-SST extensions can be used, as they have not been simulated for any site in this Hemisphere. The Southern Hemisphere comparison is performed using telescopes present within the **STD**, the only configuration present in every simulated Southern site.

In order to assess the best layout candidate (between those of equal cost), all array configurations present at a certain simulated site can be used. Both “Leoncito” and “Aar” sites are the best suited for this purpose, as all different layout configurations have been simulated (**STD**, **SCMST** and 4m-SSTs extensions) in these locations.

The chosen candidate layouts to assess the effect of the construction site on performance are the sub-arrays “2N” for the CTA-N, consisting of 4 LSTs and 15 MSTs and “2A” for the CTA-S, composed by 4 LSTs, 24 MSTs and 35 7m-SSTs, both shown in Fig. 3.14.

Due to the cost of the 7m-SST, “2A” layout turned out to be considered unrealistic. For this reason other options using 4m-SSTs are more likely to be built. Using **STD** and **SCSST** configurations, new candidate layouts have been proposed for CTA Southern site: “2Q” made up of 4 LSTs, 24 MSTs and 72 4m-SC-SSTs and the “2Z” layout, composed of 4 LSTs, 24 MSTs and 68 4m-SC-SSTs. Both candidates are shown in Fig. 3.15. “2Z” uses a more symmetric distribution of MSTs while “2Q” uses the same distribution of MSTs than “2A” candidate array.

Concerning the Northern site layout, alternative telescope positions are pro-

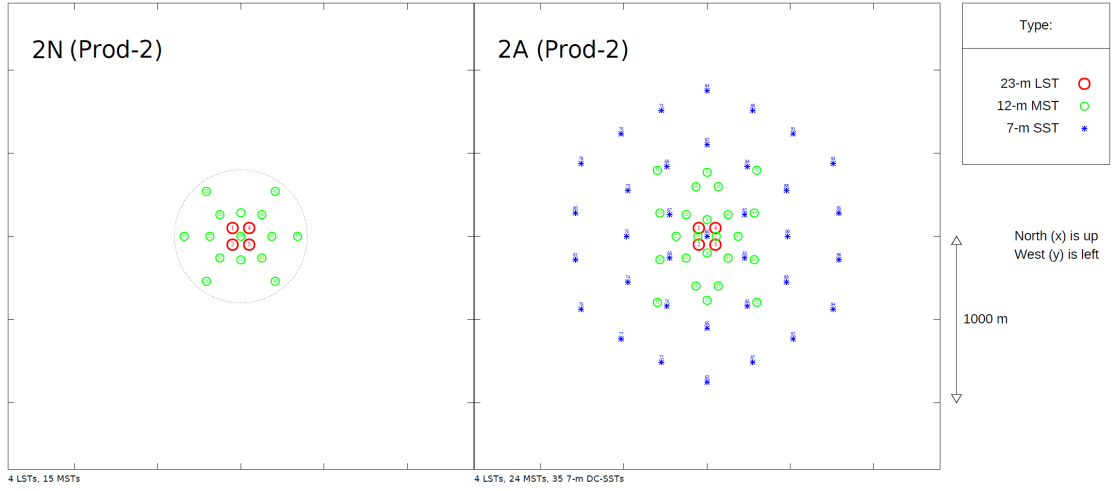


Figure 3.14: Proposed layouts to compare CTA performance over different construction sites. *Left:* The Northern Hemisphere candidate “2N” layout, made up of 4 LSTs and 15 MSTs. *Right:* The Southern Hemisphere candidate layout “2A” composed of 4 LSTs, 24 MSTs and 35 7m-SSTs.

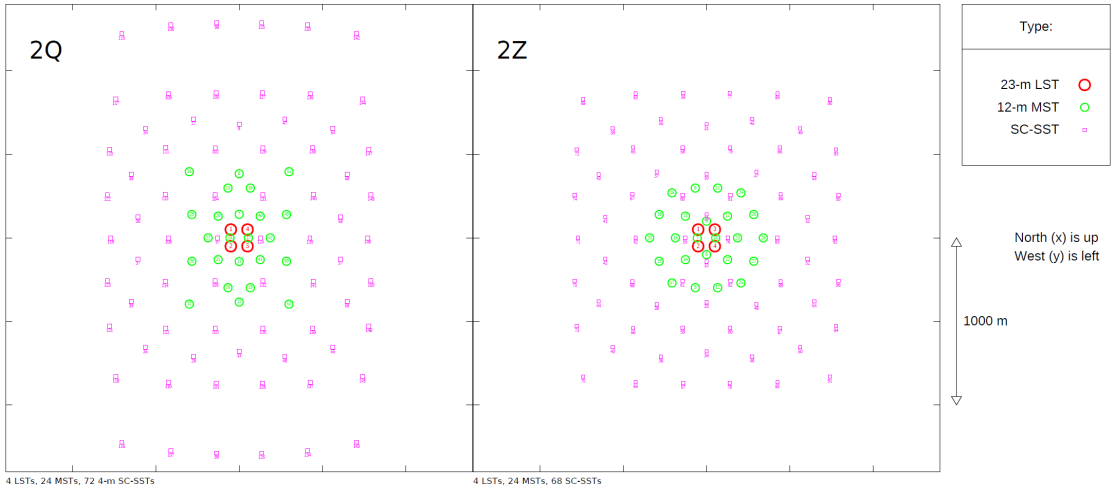


Figure 3.15: Proposed layouts for the Southern Hemisphere site, using 4m-SC-SSTs. *Left:* “2Q” candidate layout, made up of 4 LSTs, 24 MSTs and 72 4m-SC-SSTs. *Right:* candidate layout “2Z” composed of 4 LSTs, 24 MSTs and 68 4m-SC-SSTs.

3. SENSITIVITY STUDIES FOR THE CTA

posed in Sec. 3.6.4 taking into account the terrain orography of the candidate site of the Observatorio Roque de los Muchachos (ORM).

3.4.2.2 Layouts with mixed MST types

At the time this work was written, the CTA Southern Hemisphere layout was planned to contain both DC and SC-MSTs telescopes. A set of 6 different layouts have been proposed (shown in Fig. 3.16) to test the best approach to integrate the two different types of MSTs into the array. Several construction approaches were tested in this work: the *interleaved* placement (see “2I” and “2KD” subarrays), with both telescope types covering the central region of the array and the *halo* option (“2H” and “2KC”), placing MSTs in the center of the array and SC-MSTs encircling them. The spacing efficiency is also tested by comparing a compact distribution of telescopes (“2H” and “2I”) against a graded one (“2KC” and “2KD”). In order to contrast MST and SC-MST absolute performance, equivalent layouts with pure DC-MSTs (“2KA”) or SC-MSTs (“2KB”) are also proposed.

All proposed layouts have equal number of LSTs (4) and SSTs (72) with identical positions and very similar number of MSTs and SC-MSTs, ranging between 24 and 26 of each type (in the case of “2KA” and “2KB” layouts, 50 DC/SC-MSTs). Note these layouts have a higher number of telescopes than the currently expected final layout of the CTA, so these will only be used for performance comparisons.

3.5 CTA performance

In this section, the performance of the future CTA observatory is evaluated, comparing different candidate layouts using Prod-2 simulated data. Most of the curves shown in this chapter are among the following IRFs:

- Differential sensitivity: Is defined as the minimum flux emitted by a steady γ -ray source detectable by the observatory in a certain period of observation time, generally 50 hours, as a function of the estimated energy.
- Angular resolution: Represents the quality of the direction reconstruction. It corresponds to the angular distance which contains the 68% of the recon-

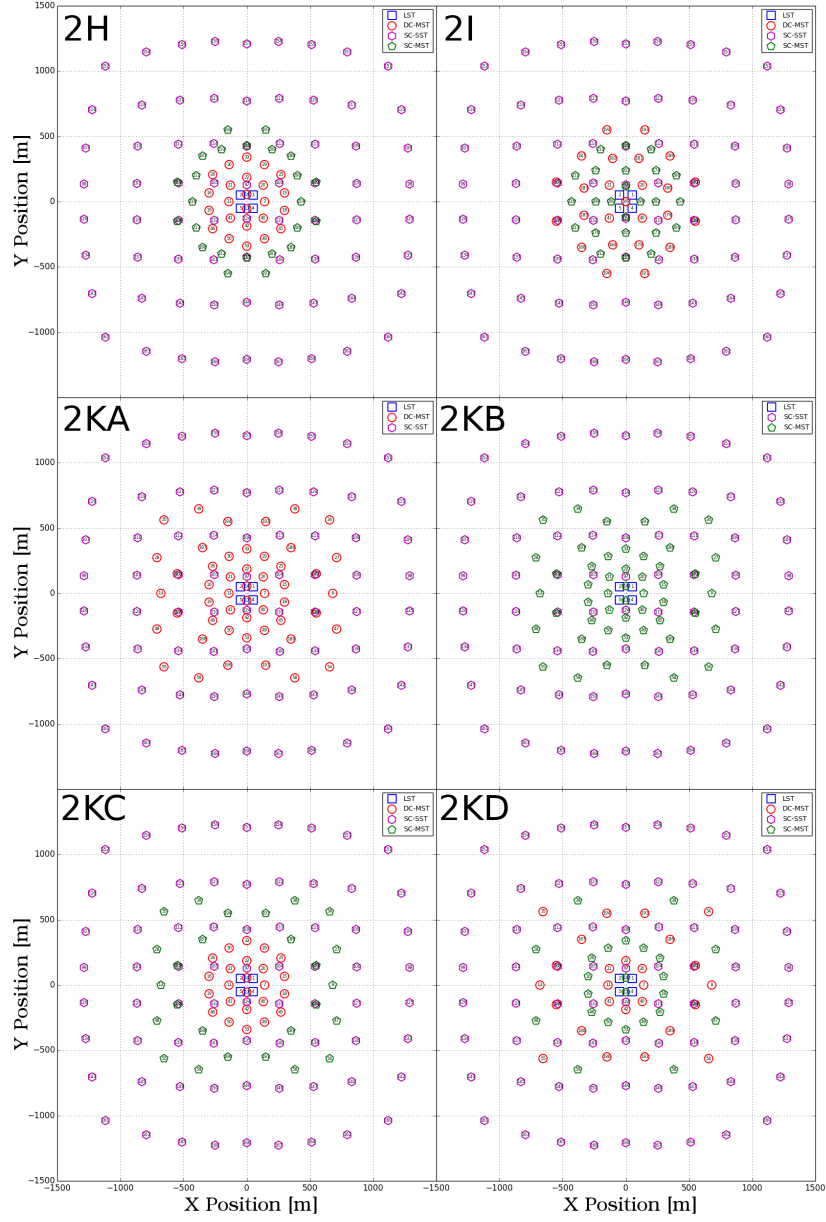


Figure 3.16: Proposed layouts to assess the impact of mixed MST type layouts on the CTA sensitivity by applying different approaches to their location and spacing. Blue squares correspond to LSTs, red circles to DC-MSTs, pink hexagons to SC-SSTs and green pentagons to SC-MSTs. Telescope placement is tested using the *halo* (“2H” and “2KC”) and the *interleaved* (“2I” and “2KD”) approaches. Telescope spacing efficiency is also explored by comparing a *compact* (“2H” and “2I”) with a *graded* (“2I” and “2KD”) distribution.

3. SENSITIVITY STUDIES FOR THE CTA

structed γ -rays relative to their true direction, directly related to the size of the PSF of the instrument as a function of the estimated energy

- Energy resolution: Accounts for the quality of the energy reconstruction. It is defined as the half width of the $\pm 34\%$ interval around the normalized most probable reconstructed energy.

All results shown in this chapter were produced using the MARS based analysis, described in section 3.3.3. As shown in Fig. 3.17, results are consistent with other analysis chains developed by other groups within the CTA Consortium (see Sec. 3.3.2).

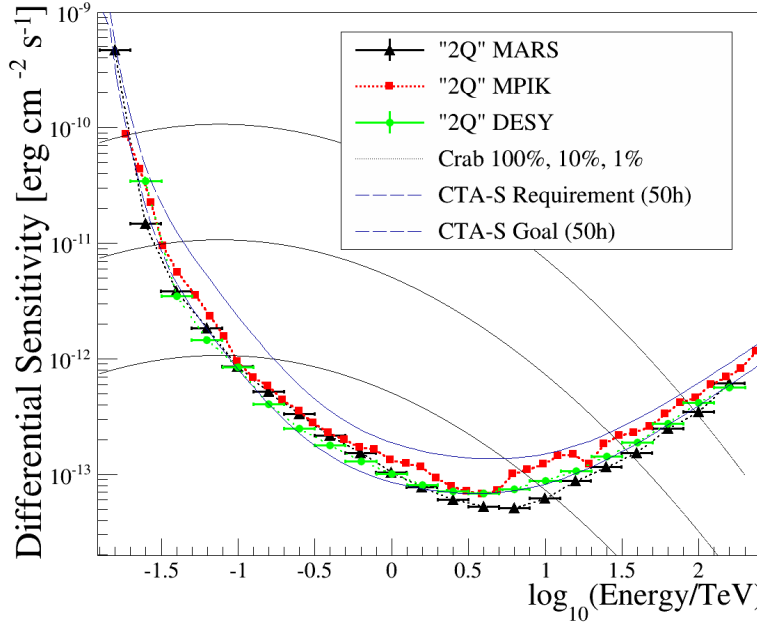


Figure 3.17: Differential sensitivity for the CTA-S “2Q” candidate (50 hours of observation, N/S pointing average) calculated with 3 alternative analysis chains: *Black*: MARS analysis, same as all results within this chapter. *Red*: *Baseline* analysis, performed by K. Bernlöhr at MPIK. *Green*: *Evndisplay* analysis, performed by G. Maier at DESY.

The following sections are devoted to show the MC analysis results describing CTA performance. In section 3.5.1, both CTA-N and CTA-S plausible candidates are analyzed in detail, comparing their IRFs with the requirements imposed by

the CTA Consortium. Section 3.5.2 describes the required IRFs optimizations performed to fulfill these requirements. This block ends with the attained results concerning the layout design study of mixed MST types in Sec. 3.5.3.

3.5.1 Plausible candidates performance

In this section the CTA capabilities are analyzed and discussed in detail. Here, the layouts considered will be the “2N” candidate array (4 LSTs and 15 MSTs) for the CTA-N layout simulated at Tenerife and the “2Q” layout (4 LSTs, 24 MSTs and 72 SC-SSTs) for the CTA-S simulated at Namibia. The choice of sites is not related in any sense with the final construction location, and is based on availability of the simulated data files (as some productions in certain sites had known bugs) and the number of cross-checks with alternative analysis chains. The choice of layouts is also relatively arbitrary, although these two layouts are considered to be good representatives of the future CTA baseline by the collaboration.

All these results correspond to the second large-scale production (Prod-2), introduced in section 3.4.2. The IRF optimization described in section 3.5.2 was used in order to fulfill these requirements in the whole energy range.

On-axis performance: The minimum detectable flux from a steady γ -ray source as a function of the energy (differential sensitivity) for both CTA sites is shown in Fig. 3.18 for 0.5, 5 and 50 hours of observation time. These curves show the sensitivity attained by averaging 2 different simulated pointing directions (north and south pointing at 20° of zenith angle). Both “2Q” and “2N” layouts well fulfill sensitivity requirements in the whole energy range, surpassing the design goals in certain regions of the spectrum. Both arrays will reach 10 mCrab sensitivity in ~ 5 hours, improving by an order of magnitude the required observation time to detect such γ -ray fluxes. It must be noted that from ~ 100 GeV up to 10 TeV, the main background contribution comes from electron induced showers. No methods for γ - e^- separation were implemented in the analysis, as the background suppression RF is solely trained with γ -ray and hadron simulated showers.

As shown in Fig. 3.19, the CTA won’t just bring an improvement in sensitivity with respect to the current instruments: it will also improve the angular and

3. SENSITIVITY STUDIES FOR THE CTA

energy resolution, shrinking by a factor ~ 3 the radius of the PSF of the current instruments, resolving structures larger than ~ 2 arc minutes, becoming the γ -ray detector with the best resolution ever constructed. Energy resolution will also improve significantly, reaching values close to the 5% in relative energy resolution. This is crucial to observe spectral features such as emission lines or to characterize the waggles appearing in the spectrum of blazars at high redshift caused by the EBL absorption.

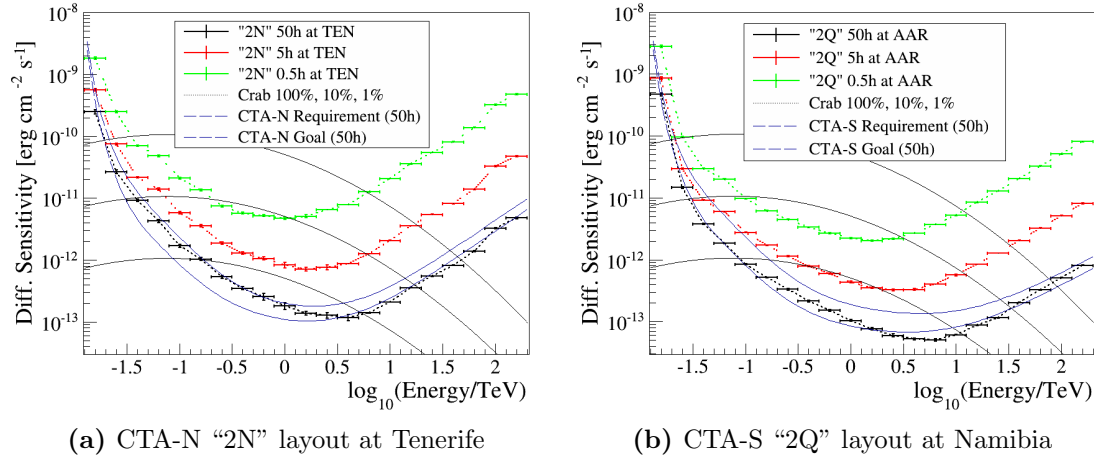


Figure 3.18: Differential sensitivity for 50, 5 and 0.5 hours for CTA-N and CTA-S candidates compared with the requirement in 50h (north/south pointing average): (*Left*) CTA-N layout candidate “2N” simulated at Tenerife. (*Right*) CTA-S layout candidate “2Q” simulated at Namibia.

To understand the overall performance of the CTA, it is required to recognize the contribution of each telescope type at different energy ranges. Performing independent analysis of sub-layouts of each telescope type within the CTA-S candidate array, is possible to gauge each telescope input to the differential sensitivity. Fig. 3.20 shows the detached performance of each layout of individual telescope types contained in the CTA-S candidate:

- **Below 100 GeV:** The subset of 4 LSTs dominate sensitivity below 100 GeV. Although this range may seem small, a huge amount of physics cases will be concentrated in this region, and cross-calibration with other γ -ray detectors, such as *Fermi*-LAT is desired.

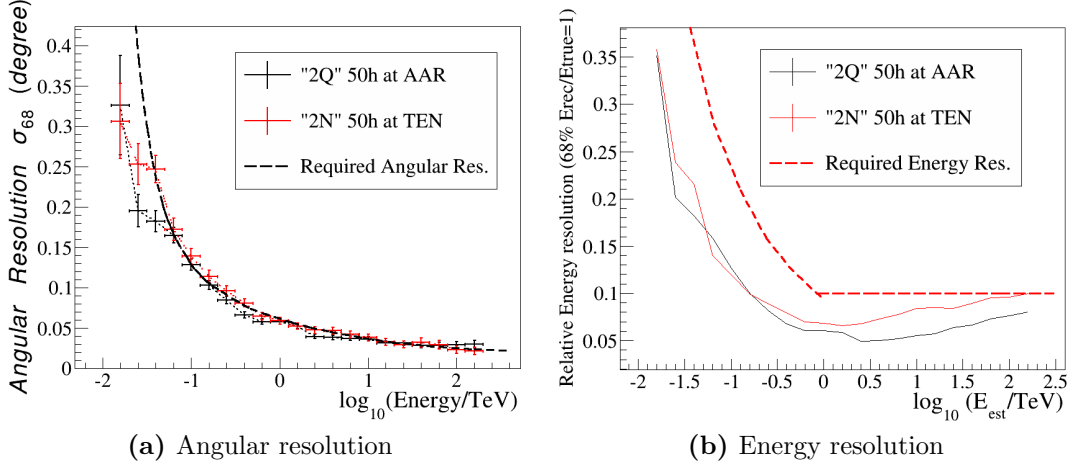


Figure 3.19: Angular and energy resolution for both CTA-N and CTA-S candidates with respect to the reconstructed energy: (*Left*) The angular resolution as a function of the reconstructed energy, defined as the angle containing the 68% of the reconstructed gamma-rays, relative to the true direction. (*Right*) The energy resolution defined as the half width of the $\pm 34\%$ interval around the normalized most probable reconstructed energy.

- **Between 100 GeV to 5 TeV:** The layout of 24 MSTs leads the instrument performance in the core energy range, above 100 GeV up to 5 TeV. Their contribution to the lower energies is also significant, as they are able to reconstruct events of ~ 60 GeV (improving sensitivity by $\sim 20\%$). They also improve gamma-hadron separation of events observed by the LSTs, as they detect possible Π_0 sub-showers or muons originated in hadronic EASs. They also contribute to higher energies up to ~ 12 TeV.
- **Above 5 TeV:** The array of 72 SC-SSTs, covering an area of ~ 4.5 Km², dominates sensitivity above 5 TeV. Different types of SSTs are expected to be spread over large distances from the center of the array, with sizes ranging between the initially designed 7m SSTs, down to 1m SSTs. Note these telescopes are only planned to be constructed in the Southern hemisphere site.

Off-axis performance: Off-axis capabilities are crucial for a significant amount of the CTA key science projects. Performing the galactic and extragalactic surveys

3. SENSITIVITY STUDIES FOR THE CTA

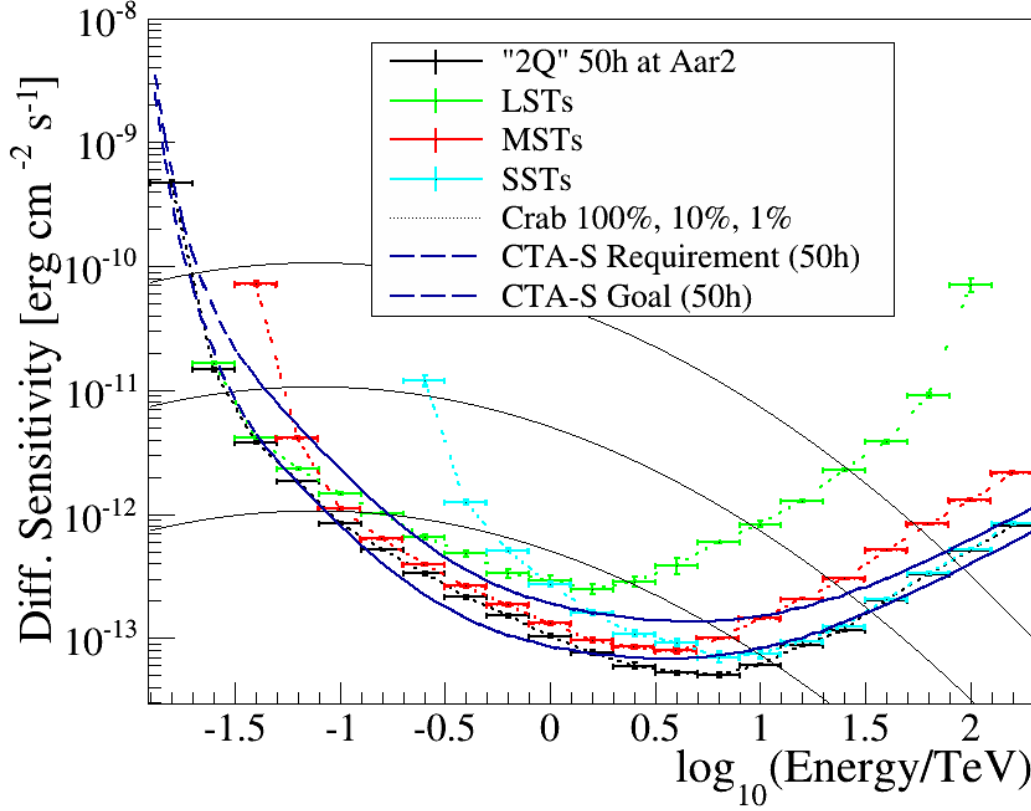


Figure 3.20: Different telescope types contributions to the over-all differential sensitivity of CTA-S candidate layout “2Q”, simulated at the Namibian site (50 hours, north/south pointing average). LSTs govern the low energy range below 100 GeV. MSTs dominate in the CTA core energies, up to 5 TeV, where SSTs start to monopolize performance.

or the detection of diffuse emission (for instance, emitted by hadronic interactions or DM signatures) relies on the instrument wide Field of View (FoV), and event reconstruction quality away from the camera center. To characterize the observatory performance for different off-axis angles, the differential sensitivity of a point-like source located at different distances from the center of the camera is estimated, as described in section 3.3.3.4.

Fig. 3.21 shows the differential sensitivity of the CTA-S candidate layout “2Q” for different off-axis angles, ranging between 0° (the on-axis performance) up to 5° . The main conclusions of this analysis are the following:

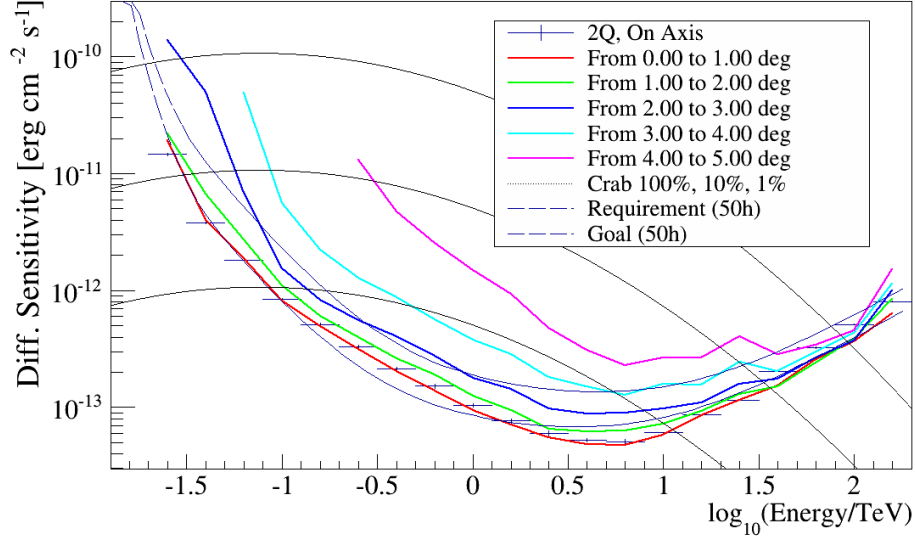


Figure 3.21: Differential sensitivity of a point-like source located at different distances from the camera center in 50 hours of observation of CTA-S candidate layout “2Q” (north/south pointing average) simulated at the Namibian site. On-axis performance is also shown for comparison. Note this analysis is strongly affected by the limitations of the generated diffuse γ -rays and hadrons statistics.

- **Below 100 GeV:** The low energy range performance, as we saw in Fig. 3.20, is dominated by the LSTs. These telescopes have a smaller field of view, coming from their parabolic reflector setup, and do not perform well at large off-axis angles. Their sensitivity does not decrease significantly at distances from the camera center smaller than 2° , but performance is drastically reduced at larger distances.
- **Between 100 GeV to 5 TeV:** In this range, the MSTs dominate sensitivity. As previously described, the modified Davies Cotton (DC) optics improves off-axis performance with respect to the parabolic one by adding a negligible spread to the photons time of arrival. In this energy range, the sensitivity drops by a factor ~ 2 on sources 3° away of the center of the camera.
- **Above 5 TeV:** As the energy increases, off-axis performance improves. This effect is related to the good off-axis performance of the Schwarzschild-Couder (SC) optics of the SC-SSTs dominating these energies, and the fact that

3. SENSITIVITY STUDIES FOR THE CTA

gammas coming from directions farther from the center of the camera can be observed at larger distances (as the CoG of their shower images fit inside the opposite side of the camera). Performance is expected to be approximately flat up to 5° away of the camera.

An analog analysis was performed for the CTA-N candidate. Fig. 3.22 shows the relative off-axis sensitivity along different energy bins of several layouts for both CTA-N and CTA-S. The Northern site shows worst off-axis performance at the low energies due to the reduced number of MSTs, but performs relatively good at higher energies, taking into account this layout does not have any SC-SSTs.

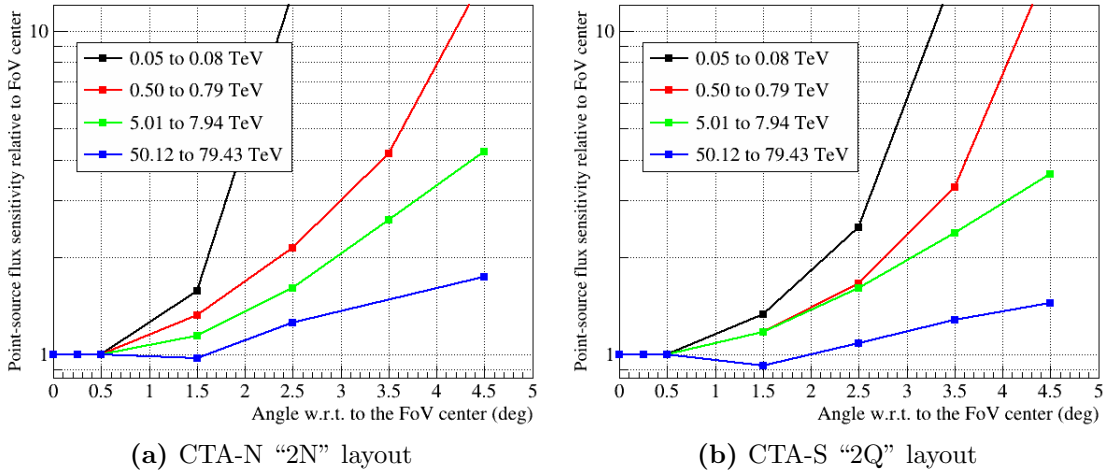


Figure 3.22: Relative off-axis sensitivity along different energy bins (50h, N/S pointing average). *left:* CTA-N “2N” layout simulated at Tenerife. *right:* CTA-S “2Q” layout simulated at Namibia, same as in Fig. 3.21.

Given the flexibility of operation of the CTA, layout subsets will be able to operate independently. The CTA Consortium opened the possibility to devote the MSTs and SSTs layout to perform several sky surveys (both galactic and extragalactic) while the central LSTs could be dedicated to other scientific cases centered in the low energy regime (between 20 GeV up to ~ 1 TeV). A dedicated analysis of the subset of MSTs and SSTs contained within the “2Q” layout was performed, in order to estimate the CTA off-axis capabilities without the contribution of the LSTs. Resulting sensitivity shows good off-axis performance over the

whole energy range, between 100 GeV up to more than 200 TeV, with an off-axis sensitivity falling no more than a factor 2 at sources 3° away from the camera center.

3.5.2 IRF optimization

As described in Sec. 3.3.3.4, the standard performance estimation optimizes parameter cuts to maximize point-like sensitivity. But different scientific cases may require to boost different capabilities of CTA performance. In particular, angular and energy resolution may be specially affected by this maximization and may not fulfill the requirements imposed by the CTA Consortium.

Different approaches to improve specific IRFs were performed using candidate “2Q” layout simulated at “Aar” site, testing CTA analysis flexibility. *Multiplicity* turned out to be the most efficient parameter to control angular resolution at CTA highest energies, improving the angular resolution up to 60%, worsening the sensitivity by a factor 2 (see Fig. 3.23). Requiring events to be observed by more than 6 telescopes at the highest energies turned out to fulfill angular resolution and differential sensitivity requirements above 30 TeV.

Using a different approach, imposing more restrictive cuts over *hadroness* and θ^2 (decreasing cut efficiencies) resulted in an improvement of the energy resolution of up to 30%, with a significant loss in the sensitivity at the highest energies of about a factor 3.

Note the current analysis should be optimized to measure the quality of each individual reconstructed event in order to improve the *multiplicity* optimization. The current applied cut, sets a *multiplicity* value through discrete steps wasting every event observed by a fewer number of telescopes, even if some of these were accurately reconstructed. During the direction reconstruction of events with multiplicities higher than 2, a minimization is performed to find the point where the distance between shower reconstructed directions is minimum. The minimum value of the minimized function is recorded, and was tested as a global parameter to assess the quality of the direction reconstruction. Several fixed cuts over the whole energy range were applied resulting on an improvement of 20% of the angular resolution above 20 TeV with a negligible effect on sensitivity. Finding efficient

3. SENSITIVITY STUDIES FOR THE CTA

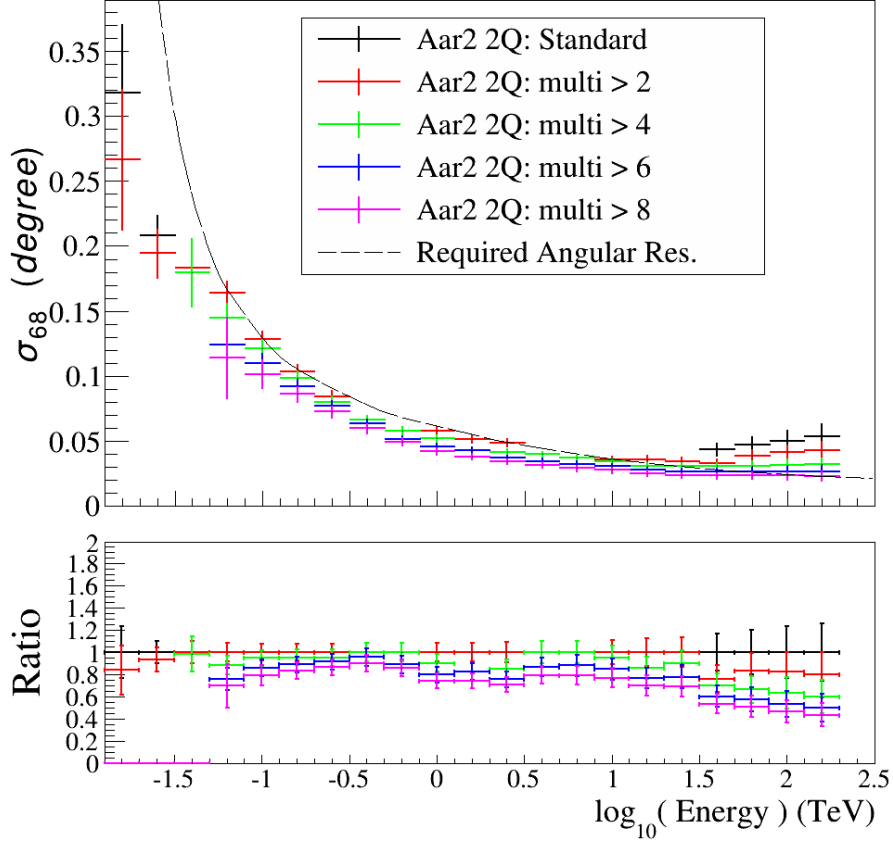


Figure 3.23: Angular resolution (corresponding to the 68% containment radius) of the “2Q” candidate array simulated at Namibian “Aar” site. Different colors show an increasing requirement in *multiplicity* reaching improvements in angular resolution at the highest energies of up to 60%. Note with this analysis we are not maximizing sensitivity anymore, and is reduced by a factor ~ 2 .

cuts of this value as a function of the energy could lead to improved angular resolution and therefore enhanced sensitivity, and may be a feasible alternative to the *multiplicity* cuts.

3.5.3 Layout design study of mixed MST types

By evaluating the layouts proposed in Sec. 3.4.2.2, important conclusions are reached concerning the most efficient approach to distribute the extension of SC-MSTs with respect to the standard array. As the analysis used in this work has been tuned for classical PMT cameras, it may not be as efficient for cameras as densely pixelated as the ones mounted in these telescopes, so these results should be considered conservative in terms of SC-MST performance. Note these layouts are not realistic scenarios of construction, and are only simulated to compare MST types performance. In Sec. 3.5.3.1, SC-MST integration in the analysis is discussed and evaluated in comparison with MST performance and in Sec. 3.5.3.2 different mixed telescope allocations and spacing are tested reaching important conclusions regarding the CTA layout construction. Lastly, Section 3.5.3.3 compares the off-axis performances of DC/SC-MST layouts.

3.5.3.1 SC-MST telescope performance

The Schwarzschild-Couder medium-sized telescopes (SC-MST) is the first ever designed IACT with a double mirror optics setup, allowing an improved angular resolution and a decrease of the focal plane scale, making possible the use of more compact cameras. This telescope is designed to hold a 11328 SiPM camera covering up to 9° of sky. Such densely populated camera allows unprecedented detail in the extracted information of the development of the cascade, although current analysis methods may still not be tuned for such capabilities. In fact, the only CTA MC analysis properly tuned to correctly assess SC-MST telescopes performance is the SLAC analysis. As described in Sec. 3.3.2, this analysis uses a parallel simplified simulation of telescopes response and cannot be cross-checked by parallel studies. For this reason, an accurately tuned analysis of the available Prod-2 SCMST configuration was required.

In order to introduce the SC-MST within the CTA MARS based analysis,

3. SENSITIVITY STUDIES FOR THE CTA

some parameters had to be tested in order to properly assess their performance. Both the trace integration time window and the image cleaning thresholds had to be tuned. Taking into account the form of a single photo-electron response of the SiPMs (broader than classical PMTs), a 10 ns window was chosen. Note the 2-pass trace extraction method applied (described in Sec. 3.3.3.1) shows much better performance for classical pixel sizes compared with the SC-MST camera. As shown in Fig. 3.24, in the case of the MSTs NSB contribution (bump at low pixel amplitudes) is reduced up to a factor 4, while in the SC-MSTs NSB contribution is barely reduced, up to a $\sim 20\%$. The reason may be related with the wider width of the window of integration, extracting more NSB regardless of the position (in time) of the extraction window, or with the lower ratio of signal pixels with respect to background pixels of the camera. The SC optics should not be the cause, taking into account the 2-pass trace extraction is also efficient for SC-SST telescopes, reducing NSB contribution by a factor 3.

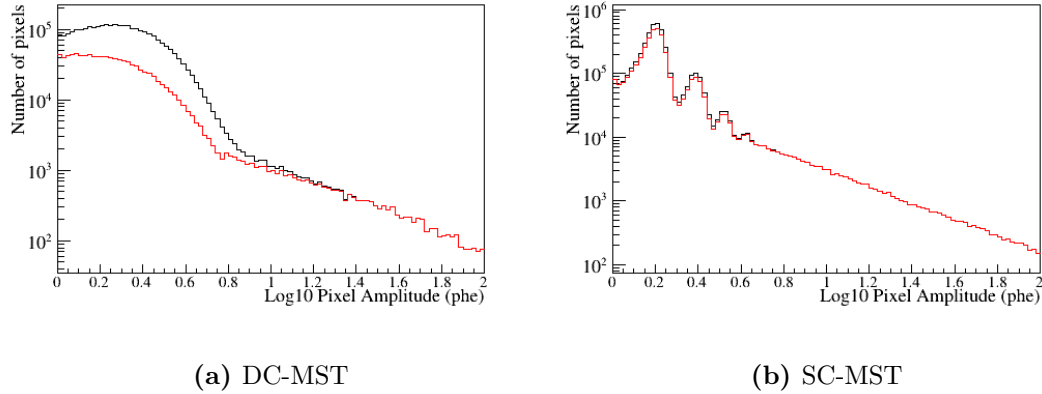


Figure 3.24: Histograms of the logarithm of the pixel amplitude (extracted charge from individual pixels) of two different MST types. *Black lines* show the extracted signal in the first pass, and *red lines* show it after the second pass, using information from the time of arrival. Note SC-MSTs highly pixelated digital camera is not tuned for the 2-pass trace integration method.

Regarding cleaning thresholds, as accidental runs were not available, several options were tested (shown in Fig. 3.25) resulting in the selection of the 2-level cleaning thresholds of 2.5/1.25 phe. Note this cleaning algorithm is not ideal for such densely pixelated cameras either, as the probability of creating higher

number of islands significantly increases with cameras with 5 times more pixels and geometries with higher number of neighbours. An alternative cleaning method, the “sum cleaning” introduced in [144], was implemented in the analysis but further tuning needs to be applied for competitive results.

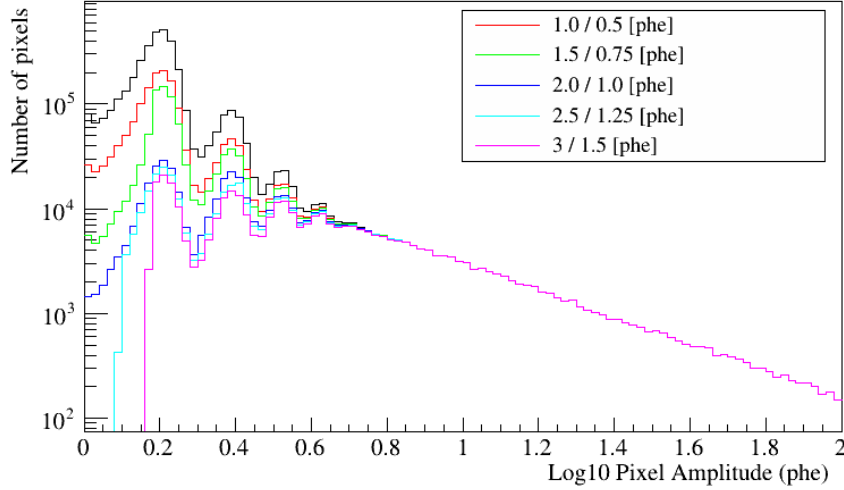


Figure 3.25: Histogram of the logarithm of the pixel amplitude (extracted charge from individual pixels) of the SC-MST for different image cleaning thresholds. Note the low amplitude peak saturates with a cleaning above 2 phe.

To test SC-MST performance and their inclusion within the CTA MARS analysis, 2 layouts were studied and compared: Both arrays have equal telescope distribution, but different MST telescope type. The first layout is the 50 DC-MSTs contained within layout “2KA” and the second one is the 50 SC-MSTs present in the candidate array “2KB”. Comparing pure DC/SC MST layouts will allow us to compare each telescope performance, and to find out their strong and lacking characteristics. Figures 3.26a and 3.26b show the differential sensitivity and angular resolution of these layouts, composed by 50 DC-MSTs and 50 SC-MSTs respectively, simulated at the Leoncito site (in 50 hours of observation, averaged between $\pm 20^\circ$ of zenith angle).

Although SC-MSTs do not help much below ≈ 100 GeV due to their smaller reflectors area, they show good performance above that point, improving the sensitivity of the core energies of CTA up to a 50% with respect to the DC-MSTs.

3. SENSITIVITY STUDIES FOR THE CTA

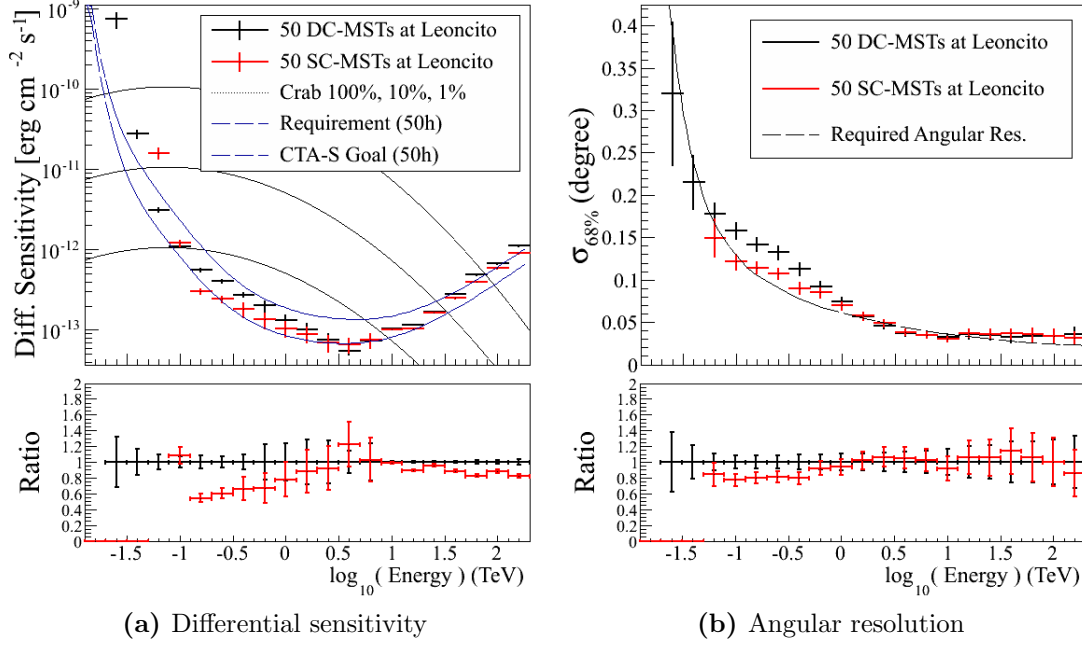


Figure 3.26: Comparison of the differential sensitivity in 50 hours (*Left*) and the angular resolution (*Right*) between 2 layouts of equal telescope distribution, one composed by 50 DC-MSTs and the other one by 50 SC-MSTs. Both layouts were simulated at the Leoncito site, with $\pm 20^\circ$ of zenith angle. Comparison shows SC-MSTs outperform DC-MSTs in the core energies of CTA, mainly due to their improved angular resolution.

Comparing θ^2 cut efficiencies and background rates shows that this improvement comes mainly from the enhanced angular resolution of these telescopes (due to the higher pixelization of the shower image), allowing a greater background rejection.

The next reasonable step is to compare the MST types performance under a more reasonable scenario: with full layouts composed by the mixture of LSTs, MSTs and SSTs. For this reason layouts “2KA” and “2KB” were analyzed. Results, in Fig. 3.27, show the loss in the low energy performance caused by the smaller reflecting area of SC-MSTs is compensated with their improved reconstruction in the core energies. Note these results rely on an analysis that is far from ideal, and has been mainly tuned for classical IACTs, so they should be considered as very conservative regarding SC-MST performance.

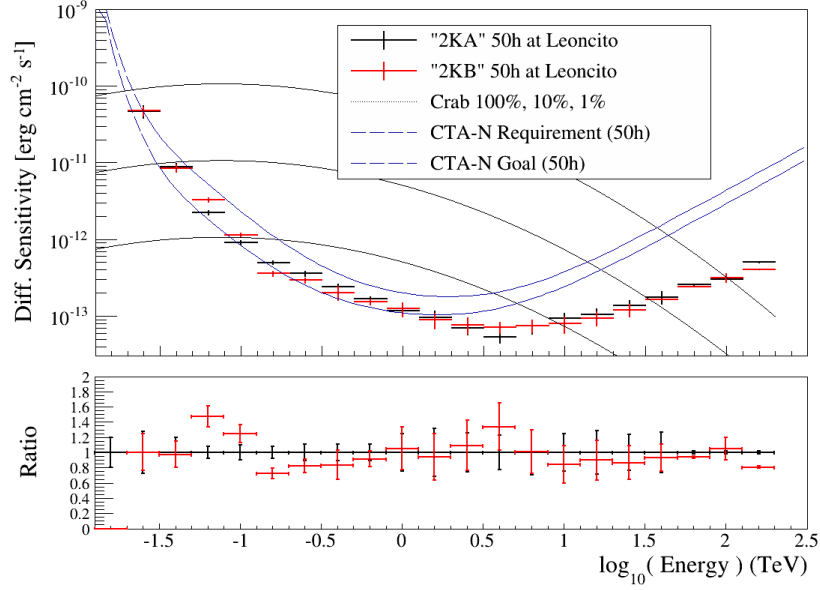


Figure 3.27: Comparison of the differential sensitivity of the “2KA” and “2KB” layouts (see Fig. 3.16) of N+S pointing average performance in 50 hours, simulated in the Leoncito site. These layouts correspond to the same telescope distributions of DC-MSTs and SC-MSTs shown in Fig. 3.26, with the addition of 4 central LSTs and 72 SC-SSTs. DC-MSTs show better low energy performance while the SC-MSTs boost the sensitivity within the CTA core energies.

3.5.3.2 Telescope spacing and distribution

Comparing the attained performance of the different proposed mixed MST type layouts, shown in Fig. 3.16, will guide us on the most efficient approach for SC-MSTs construction. These are expected to be constructed later than the DC-MSTs, so logistically the simplest approach would be to build SC-MSTs around the already constructed ones (*halo* approach). But the possibility of an improved performance by interleaving the location of both telescope types is still open. Additionally, the optimum separation of MSTs is still to be decided. Its clear larger inter-telescope distances would improve the effective area for larger energies, while closer telescopes improve the low energy coverage.

3. SENSITIVITY STUDIES FOR THE CTA

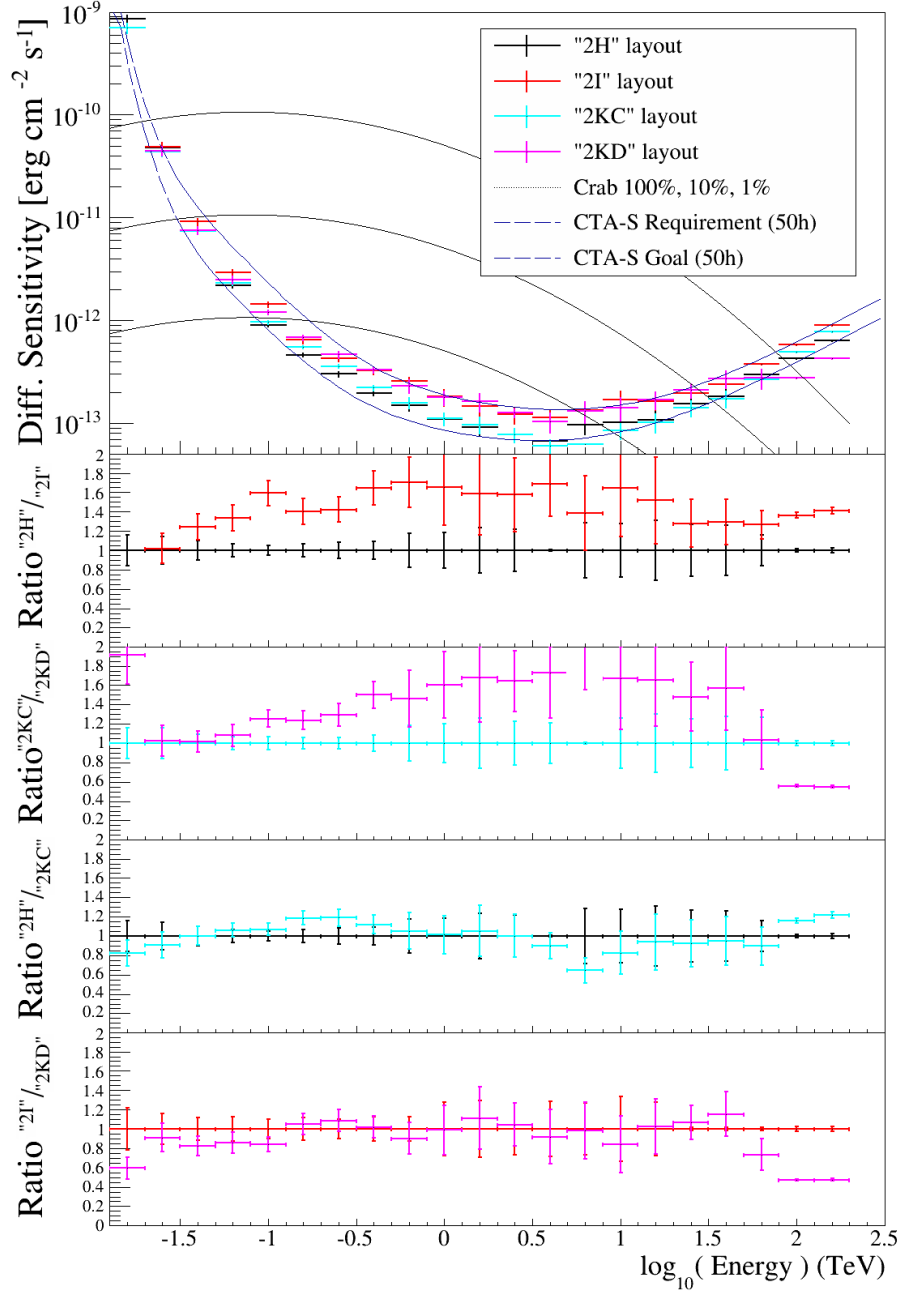


Figure 3.28: Differential sensitivity of mixed DC/SC-MST layouts (shown in Fig. 3.16) in 50 hours of observation (average of north and south pointing directions) simulated at the Prod-2 Leoncito site. First two ratio plots compare layouts following the halo approach (“2H” or “2KC”) with the interleaved option (“2I” or “2KD”), showing halo approach improves the core energy sensitivity by nearly a factor 2. Comparing compact (“2H” or “2I”) against graded arrays (“2KC” or “2KD”) show no clear preferred option.

After applying all changes to the analysis to integrate SC-MSTs, all proposed layouts were analyzed. Very similar statistics were used for the different training samples (not identical, taking into account every layout had different trigger efficiencies). As previously described, SC-MST analysis is far from being ideal, as some steps in the analysis (mainly image cleaning and signal extraction) should be improved for highly pixelated cameras.

Regarding the relative position of the different MST types in the layout, the halo approach shows clear improvement over the interleaved option. As shown in Fig. 3.28, comparison between “2H” over “2I” and “2KC” over “2KD” leaves no doubt about the sensitivity improvement of the halo approach, up to a factor 2 in sensitivity. The reason behind this improvement is the differences between each telescope characteristics. DC-MSTs have larger reflectors improving low energy events reconstruction while the SC-MSTs, with a reduced area of reflectors, have improved angular resolution and off-axis performance (shown in section 3.5.3.3). Separating the telescopes of the same type, increases the amount of events reconstructed using both types. The effect of each individual SC-MST improved direction reconstruction is weakened if stereo events are observed by telescopes of different type (with worst angular resolution), reducing the potential quality of the global direction reconstruction, as seen in Fig. 3.29. Similarly, low energy events multiplicities will be reduced, as events detectable by MSTs and not by SC-MSTs will be significantly worst reconstructed. The halo approach maximizes the amount of low energy events well reconstructed between LSTs and MSTs, and also increases the amount of higher energy events reconstructed in stereo by SC-MSTs.

Determining the most efficient telescope spacing may be harder, at least with the attained results. Comparing the compact layouts (“2H” and “2I”) against the graded ones (“2KC” and “2KD”) shows no clear preferred option. Observing the layouts using the *halo* approach, “2H” over “2KC” sensitivity ratio, a hint of improvement in the low energy range is seen by placing these telescopes closer to the center of the array, although the effect is balanced with a decreasing performance over higher energies.

3. SENSITIVITY STUDIES FOR THE CTA

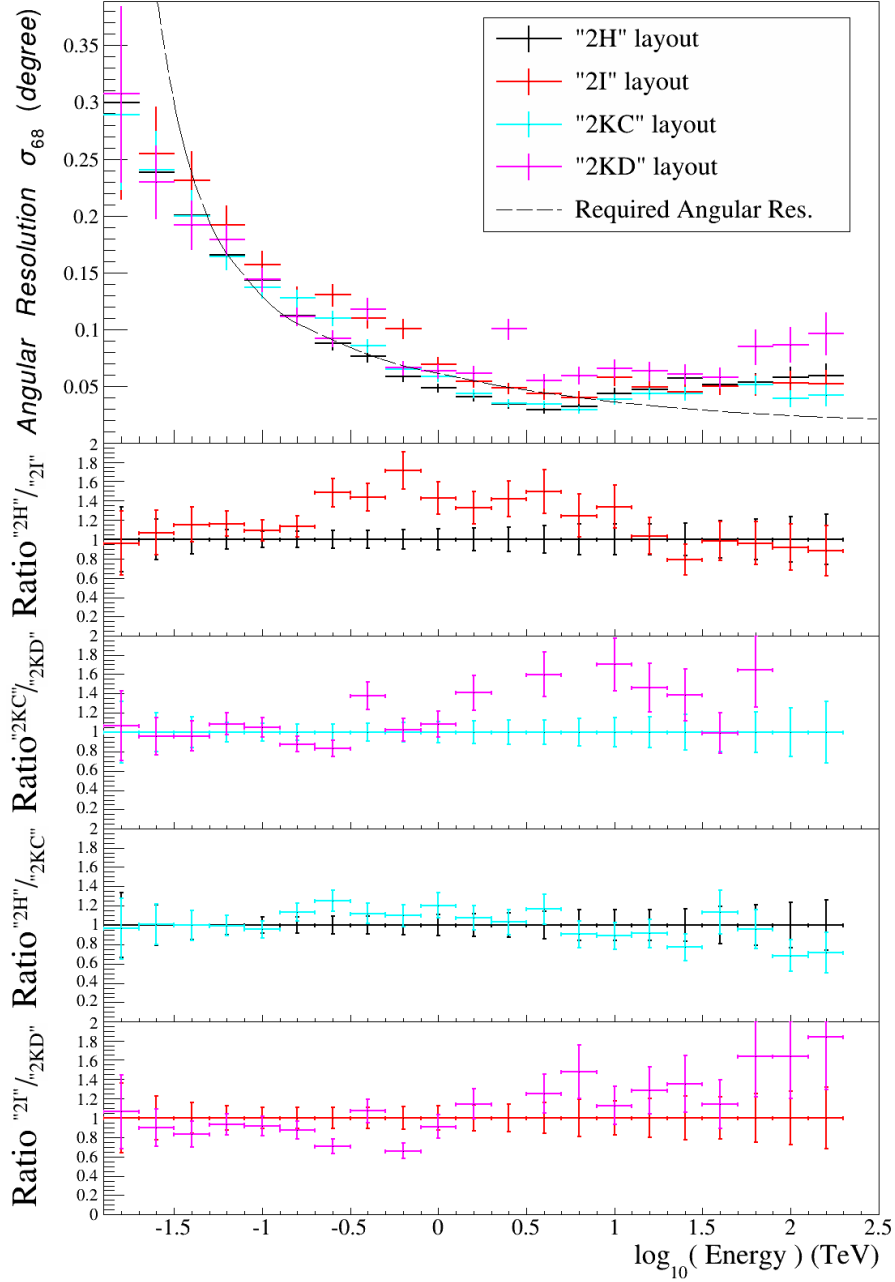


Figure 3.29: Angular resolution (68% containment) of mixed DC/SC-MST layouts (shown in Fig. 3.16) in 50 hours of observation (average of north and south pointing directions) simulated at the Prod-2 Leoncito site. First two ratio plots compare layouts following the halo approach (“2H” or “2KC”) with the interleaved option (“2I” or “2KD”), showing halo approach also improves angular resolution ($\approx 60\%$). Comparing compact (“2H” or “2I”) against graded arrays (“2KC” or “2KD”) show no clear preferred option.

It must be noted the analysis performed by the author of this work was the only one showing good performance of the SC-MST telescope, mainly caused by their higher camera pixelization. Cross-checks should be provided with other analysis chains, which should be carefully tuned to make use of their specific properties in order to perform a fair comparison between DC/SC-MSTs.

3.5.3.3 Off-axis performance

As described in section 3.5.1, off-axis performance is crucial for some of the CTA scientific objectives. In this section, the off-axis performance of different MST types is compared, assessing the impact of the SC-MST extension to the FoV.

To compare DC-MSTs with SC-MSTs, the off-axis capabilities of “2KA” and “2KB” layouts were analyzed. Of course, these layouts are not realistic due to the high number of MSTs, but they are perfect to compare their MSTs performance. In addition, the off-axis performance of “2KC” layout was also tested, to estimate the effect of the addition of SC-MSTs to the standard CTA-S layout.

Figure 3.30 shows off-axis sensitivity ratios for the three analyzed layouts. Interesting conclusions related to the off-axis performance of the different MST telescope types are inferred:

- Comparing “2KA” and “2KB” shows the low energy off-axis performance relies on DC-MSTs, as SC-MSTs do not help much below ~ 300 GeV. This result is consistent to the one shown in Fig. 3.26. Regarding higher energies, SC-MSTs outperform off-axis performance of DC-MSTs increasing the FoV by a 25% in radius.
- Comparing “2KA” and “2KC” shows the central DC-MSTs are enough to attain a similar low energy off-axis performance as in “2KA”, while graded SC-MSTs improve the higher energies FoV by a 20% in radius.

Taking into account these conclusions, the most efficient layout for improving the core energy range of the CTA while augmenting its off-axis capabilities would be a mixed DC/SC-MST, following the *halo* approach: DC-MSTs would be placed in the center near the LSTs while SC-MSTs would be placed surrounding them. Note this extension of SC-MSTs would greatly affect the amount of time needed

3. SENSITIVITY STUDIES FOR THE CTA

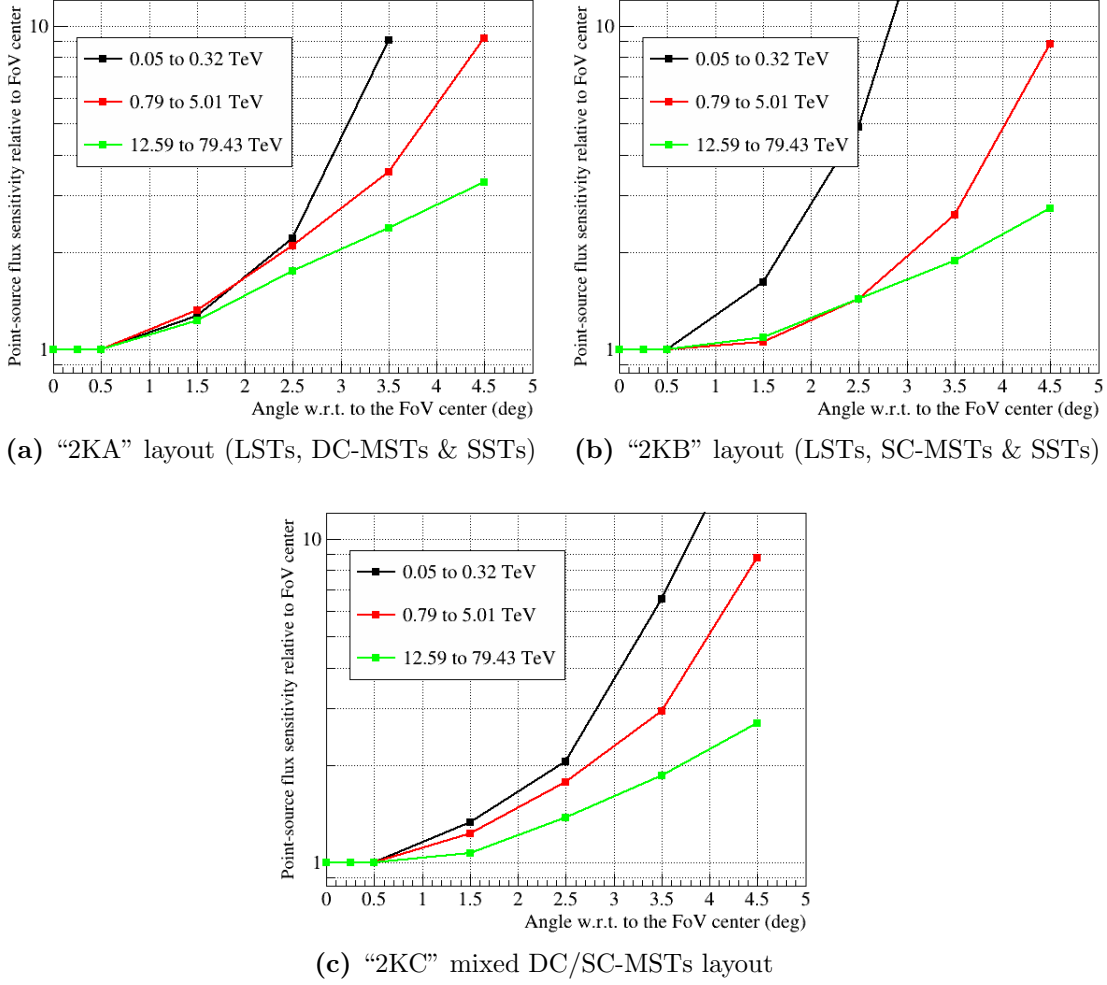


Figure 3.30: Relative off-axis sensitivity along different energy bins for 3 mixed MST types layouts (50h, N+S pointings at Leoncito site).

for the different surveys the CTA Consortium is planning to perform, and would increase the chances of serendipitous detections through the observatory lifetime.

3.6 CTA-N site selection

In addition to the estimation of the CTA performance, discussed in Sec. 3.5, one of the main objectives of the Prod-2 was the accurate evaluation of the effect of the different proposed construction candidate sites on the CTA overall sensitiv-

ity. As previously introduced, the main studied differences between the MC site configurations are the effects of the altitude, the atmospheric density profiles and their different local Geomagnetic Field (GF).

To estimate overall scientific performance of each building location, a set of parameters evaluating their capabilities were defined by the CTA Consortium, in order to perform unbiased comparisons taking into account all major effects on their performance. The following terms were defined:

- Average Annual Observation Time (AAOT): Average yearly observation time (in hours) of different construction sites. Essentially accounts for the different weather conditions, where precipitation, clouds or strong winds may not allow normal observations, therefore decreasing this number.
- Performance per Unit Time (PPuT): Geometric average along the energy of the ratio between the CTA required and calculated sensitivity, defined by the expression:

$$PPUT = \left\{ \prod_{E_0}^{E_N} \frac{S_{i,req}}{S_i} \right\}^{1/N}, \quad (3.3)$$

where $E_0 = 30 \text{ GeV}$ and $E_N = 30 \text{ TeV}$ for the case of the CTA-N and $E_N = 200 \text{ TeV}$ for the CTA-S. Accounts for the effects related to the IACT sensitivity on a given time, mainly driven by the altitude and the GF effect.

- Figure of Merit (FoM): Parameter evaluating the total potential scientific performance of a specific construction site taking into account the observatory performance and the time it will likely observe under the required conditions. It is defined as follows:

$$FoM = PPuT \cdot \left\{ \frac{AAOT}{1100 \text{ h}} \right\}^{0.7}, \quad (3.4)$$

where the 1100 h is roughly the 70% of the maximum dark sky time measured and the index 0.7 tunes the parameter not to be neither background nor signal dominated.

3. SENSITIVITY STUDIES FOR THE CTA

Therefore, the CTA sites performance is compared using PPuT values calculated from each simulated candidate site within the Prod-2 using equal telescope layouts. To do so, the performance estimated using the different analysis chains described in Sec. 3.3.2 is averaged, reducing the effect of particularities within different methods.

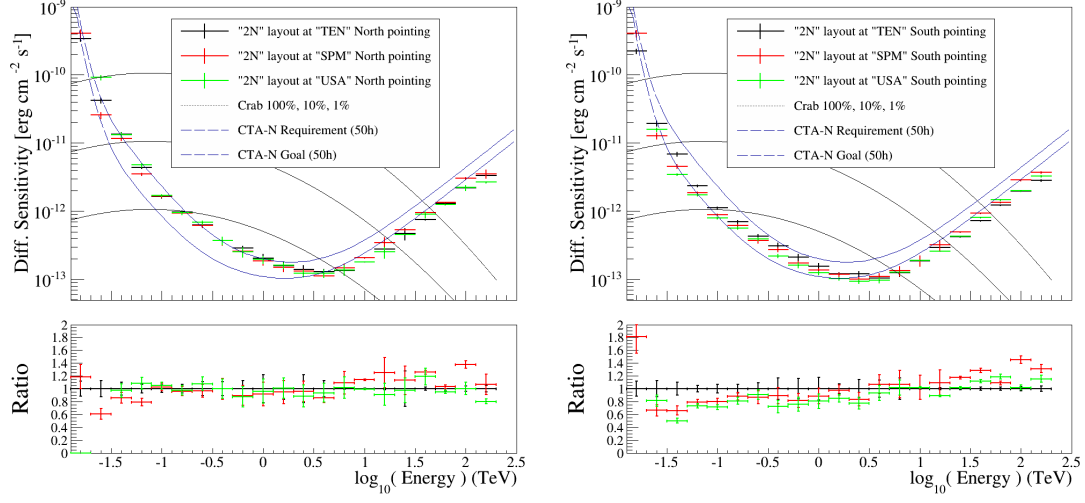
Dedicated Prod-2 simulations were generated by the Consortium for all proposed sites, in order to find their qualifications to host such a project. This section overviews the attained results on the sensitivity studies performed for the different simulated sites in the Northern Hemisphere: Tenerife (“TEN”), San Pedro Martir (“SPM”) and Arizona (“USA”).

In section 3.6.1, the performance of the different sites is compared, together with a detailed discussion about the attained results. As the effect of the geomagnetic field dominates northern site performance differences, section 3.6.2 introduces a noble method to calculate realistic values of its averaged effect. Section 3.6.3 is devoted to determine the effect of an increased NSB (30%) on the observatory performance. Finally, in section 3.6.4, orography restrictions at the *Rogue de los Muchachos* site (La Palma, Spain) are taken into account and new alternative layouts are evaluated, testing if slightly different telescope distributions would have a significant effect on final performance.

3.6.1 Site comparison

Using the standard analysis detailed in section 3.3.3, the 3 different Northern Hemisphere sites are analyzed. The same layout is used in all cases: The standard CTA-N layout “2N” (see Fig. 3.14). As previously described, north and south pointing directions (both at 20° of zenith angle) are analyzed separately, with independent direction reconstruction LUTs, energy RF and gamma-hadron separation algorithms. Obtained results are shown in Fig. 3.31a, for the north pointing and Fig. 3.31b, for the south pointing direction.

As shown in Fig. 3.31, the north pointing direction shows little differences between the simulated candidate sites performance, mainly caused by the different altitudes. On the other hand, comparing the south direction, significant differences appear below 200 GeV caused by the different GF affecting the EASs.



(a) North pointing (20° in zenith angle)

(b) South pointing (20° in zenith angle)

Figure 3.31: Differential sensitivity of the standard CTA-N “2N” layout in 50 hours of observation simulated at each site: Tenerife (“TEN”), San Pedro Martir (“SPM”) and Arizona (“USA”). Each pointing direction is shown separately to see the strong effect of the geo-magnetic field on performance: *Left*: North pointing (higher overall GF at every site). *Right*: South pointing (Lower GF with larger differences between sites).

To understand these differences, knowledge on the effect of the GF is required. As shown in Fig. 3.32, the orthogonal magnetic field affecting the simulated EASs coming from the two positions in the sky simulated within the Prod-2 is different for each site. In the North pointing, all sites are affected by a $B_{\perp} \approx 37 \mu T$, showing worst low energy performance and small differences between them. On the other hand, in the South pointing direction, Tenerife is affected by a $B_{\perp} \approx 20 \mu T$ while the two American sites by roughly $B_{\perp} \approx 10 \mu T$.

As introduced in the appendix A, the magnetic field increases the separation of generated $e^{-}e^{+}$ pairs, blurring the distribution of the Cherenkov photons arriving to the ground in the a East-West direction, producing an average loss in the Cherenkov photon density detected by IACTs at small core distances (see Fig. A.9), reducing sensitivity in the low energy range. At larger distances, it may increase the observed Cherenkov emission, improving the effective area (and therefore sensitivity) in the VHE range (above ~ 10 TeV). Another effect of the

3. SENSITIVITY STUDIES FOR THE CTA

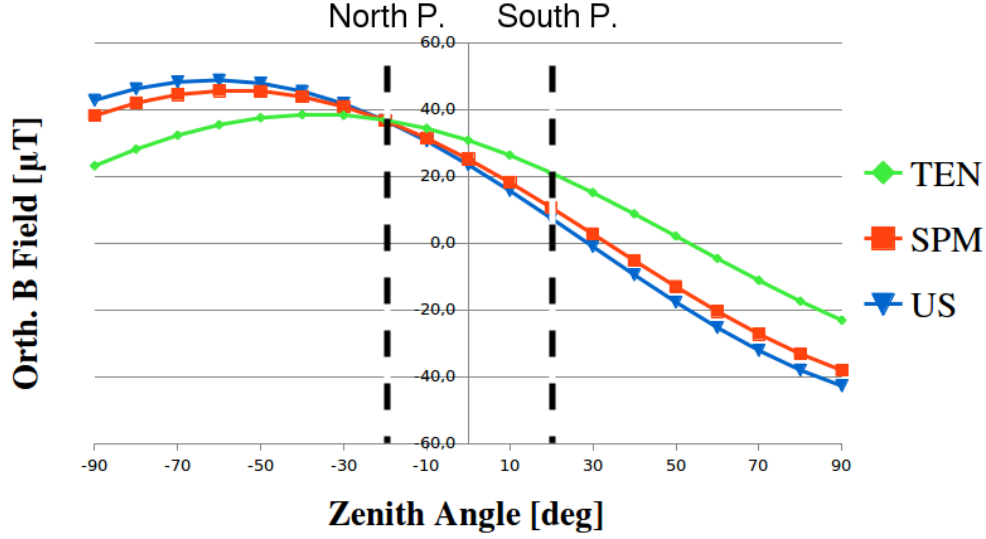


Figure 3.32: Orthogonal GF maximum intensity along the zenith angle. Dashed vertical lines show the simulated values within the Prod-2N. Comparing the averaged values of sensitivity between simulated positions ($\pm 20^\circ$ zenith, 0° (South) and 180° (North) in azimuth) may show unrealistic performance differences.

GF is the shift of the reconstructed direction of some events, as described in [149]. This effect will be corrected to a certain extent with the analysis (which currently does not take it into account).

Looking at the distribution of the orthogonal GF along the different values of the zenith angle (Fig 3.32), the region where the Spanish site has an increased level of B_\perp is roughly between -20° and $+40^\circ$, thus calculating the PPUT values using only these two simulated positions in the sky ($\pm 20^\circ$ of zenith angle) would show an artificially worst performance for Tenerife. This problem was solved by calculating the average effect of the GF per site, as described in the next section.

3.6.2 Averaged Geo-magnetic field

A strong GF field dependence was observed in the Prod-2 results, specially in the case of the CTA-N candidate sites, as the main differences in sensitivity come from the geo-magnetic field effect (see Fig. 3.31). As average sensitivities between North and South pointings are used for comparing different sites performance,

results may be misleading taking into account the limited amount of simulated pointing directions.

The effect of the GF on the development of the EAS is proportional to its orthogonal component, expressed by the following expression:

$$|\vec{B}_\perp| = \{B_z^2 \sin^2 \theta \sin^2 \phi + (B_x \cos \theta - B_z \sin \theta \cos \phi)^2 + B_x^2 \sin^2 \theta \sin^2 \phi\}^{1/2}, \quad (3.5)$$

where θ and ϕ are the *zenithal* and *azimuthal* angles of the charged particle direction (approximately parallel to the EAS development) and B_x and B_z the local components of the GF intensity, fixed by the site location. Using Eq. 3.5, we can compare the orthogonal GF intensity for the Northern sites at the simulated zenith angles. Considering observations in the azimuthal angle uniform, in Fig. 3.32 is shown the maximum orthogonal B field as a function of the zenith angle of candidate sites.

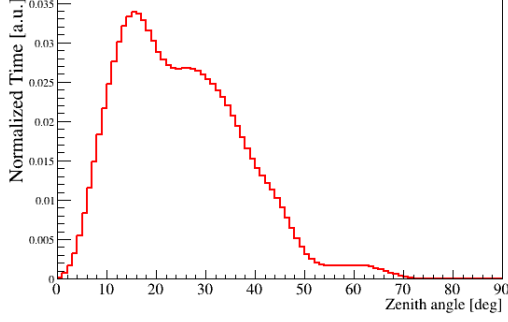
The effect of the GF depends both on the azimuth and zenith angles. As most of the performance differences come from the distinct intensities of the local magnetic fields, evaluating the average effect affecting an observatory located in each of the proposed sites weighted with the expected distribution of observations (both in azimuth and zenith angles) seems to be the best way to estimate CTA candidate layouts performance differences.

Expecting a similar distribution of observations as in other Northern Hemisphere IACT experiments, this study employed a smoothed distribution in the zenith and azimuth angles taken from the MAGIC experiment. Observations along the azimuth angle were considered uniform (which is approximately true) while a normalized smoothed zenith angle distribution of the MAGIC observations since June 2006 were used, shown in Fig. 3.33a. For each zenith angle, the average (in azimuth) orthogonal GF was calculated (shown in Fig. 3.33b) and weighted with the normalized number of hours of observation (Fig. 3.33c). Integrating that distribution results in the weighted averaged values of the GF intensity for each candidate site (Fig. 3.33d).

These values estimate the real averaged effect of the GF on the observatory performance at each CTA-N candidate site. The North American site shows the

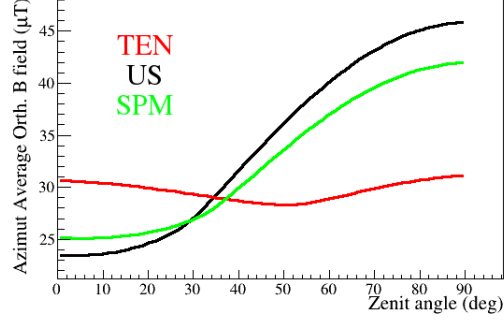
3. SENSITIVITY STUDIES FOR THE CTA

Normalized MAGIC observations since June 2005



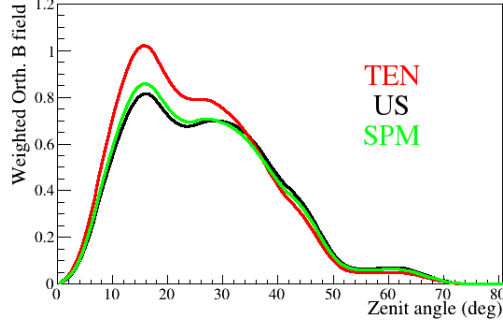
(a) MAGIC averaged observations

Average Orth. B



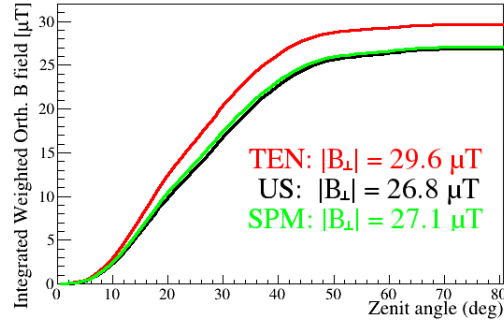
(b) Average Orth. B

Weighted Average Orth. B



(c) Weighted average orth. B

Integrated Weighted Average Orth. B



(d) Integrated average orth. B

Figure 3.33: Different steps used in the calculation of the CTA-N candidate sites averaged orthogonal GF effect. (a): Normalized smoothed zenith angle distribution of MAGIC observations since June 2006. (b): Calculated average through the azimuth angle of the orthogonal GF as a function of the zenith angle. (c): Weighted average orthogonal GF with normalized smoothed MAGIC observations. (d): Resulting integrated weighted average orthogonal GF for each CTA-N candidate site.

lowest effect with $\overline{B}_\perp = 26.8 \mu T$, followed by the Mexican site with $\overline{B}_\perp = 27.1 \mu T$ and finally the Spanish site with $\overline{B}_\perp = 29.6 \mu T$, as expected. Comparing these values with the arithmetic average of the MC simulated conditions shown in Table 3.6 (only two points at $\pm 20^\circ$ in zenith and 0° and 180° in azimuth), differences shrink more than a 60%.

Table 3.6: Average orthogonal Geomagnetic field effect on the CTA-N candidate sites

Site	Prod-2N \overline{B}_\perp ($\pm 20^\circ$) [μT]	\overline{B}_\perp (weighted with MAGIC obs.) [μT]
US	22.0	26.8
SPM	23.6	27.1
Tenerife	28.9	29.6

Averaged values of the GF affecting each CTA-N candidate site under 2 different considerations. The first set of \overline{B}_\perp values only considers simulated points in the sky within the Prod-2N, corresponding to $\pm 20^\circ$ in zenith with 0° and 180° in azimuth. The second set of values corresponds to an averaged effect taking into account the expected observations of a IACT observatory placed in the Northern Hemisphere.

These considerations were taken into account and the effect was included in the evaluation of the PPUTs of the CTA candidate sites both in the Northern and Southern Hemisphere, and further simulations using $\pm 40^\circ$ in zenith angle observations validated the process (see Table 3.7).

Table 3.7: Overall performance of simulated sites of the Northern Hemisphere

Site	PPUT				FoM
	N	S	Average	Weighted	
Tenerife	1.27 ± 0.02	1.67 ± 0.04	1.46 ± 0.03	1.45 ± 0.07	1.46 ± 0.10
US	1.30 ± 0.02	2.07 ± 0.05	1.71 ± 0.03	1.50 ± 0.07	1.40 ± 0.10
SPM	1.31 ± 0.02	1.86 ± 0.05	1.58 ± 0.04	1.58 ± 0.08	1.60 ± 0.09

Performance per unit time and figures of merit of the standard CTA-N layout simulated at each different site. Weighted PPUT values correspond to linearly extrapolated values using the \overline{B}_\perp values weighted with MAGIC observations.

3.6.3 Tenerife NSB

During the CTA site selection process (also before the NSBx3 configuration was simulated) there was no accurate estimation of the impact of the increased NSB levels affecting the Tenerife site on the observatory performance. The NSB has a significant impact in the low energy threshold of Cherenkov telescopes affecting both accidental trigger rate and the signal to noise ratio of shower images. The increased accidental trigger rate requires to increase trigger thresholds, consequently losing low energy events, while the lower signal to noise ratio produces more spurious photons, requiring increased image cleaning thresholds.

This section shows the results obtained with a simplified analysis to test the effect of an increased level of NSB, assessing the CTA high level responses (after analysis cuts). Running both CORSIKA and `sim_telarray` with the NORTH configuration at the Tenerife site, a moderate number of γ -ray showers were simulated with a 30% increase in the NSB frequency. As a full proton production needed too many resources, some considerations were taken into account in the analysis:

- To replicate the effect of the higher NSB in these shower images, increased image cleaning thresholds were applied to the simulated γ -rays, scaled with the square root of the NSB.
- Direction and energy reconstruction is performed using simulated γ -rays with the increased NSB level, showing equal performance than the default configuration.
- γ -hadron separation algorithms trained with the standard γ -rays and protons were used. This consideration implies the obtained results using this analysis should be considered very conservative, artificially increasing the effect of the NSB (as some γ -rays with the increased NSB level may be considered hadrons by the RF).
- Same *hadroness*, θ^2 and *multiplicity* cuts from the normal Tenerife analysis are applied to the two different sets of γ -rays: Those simulated with the increased level of NSB, and the standard configuration ones.

Using these considerations, the resulting ratio between γ rates after high level quality cuts are applied (ratio between the effective areas of 1/1.3 NSB level) is shown in Fig. 3.34. The main conclusions are the following:

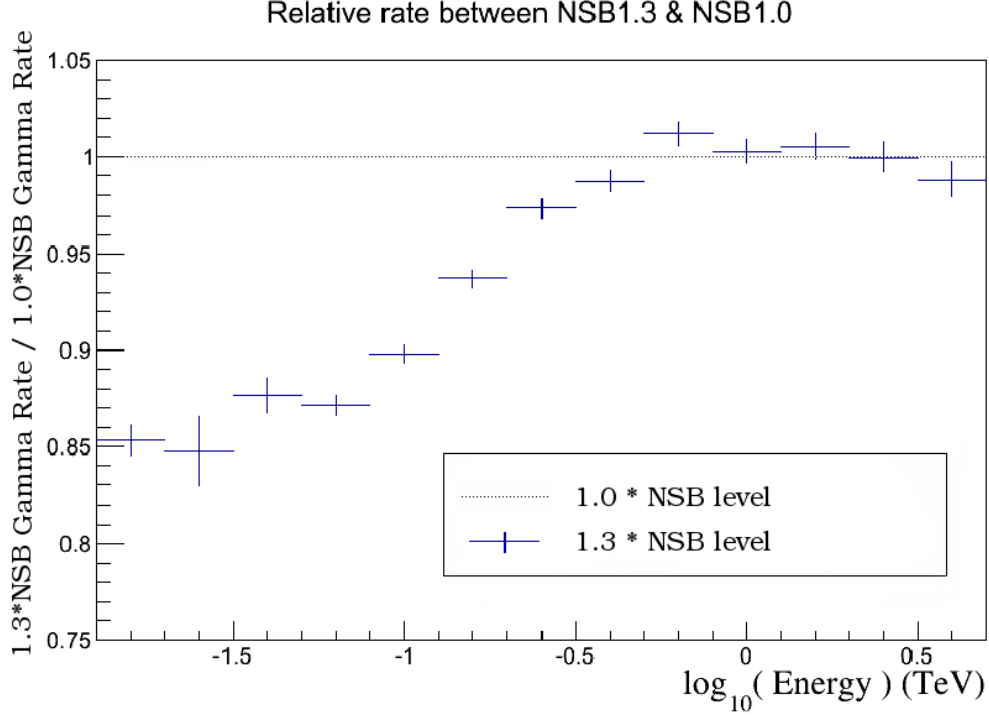


Figure 3.34: Ratio between the γ rates of different NSB levels after high level quality cuts are applied. The effect is negligible above ~ 400 GeV.

- Below 100 GeV there is an effective area loss of $\sim 10 - 15\%$, although the effect in sensitivity would not be as significant, as it is imposed by the 5% of background condition (also reduced by the increased image cleaning threshold).
- Between ~ 100 and 400 GeV, effective area loss becomes $\sim 5\%$, leading to a loss in sensitivity of $\sim 2\%$.
- Above 400 GeV, the effect of an increase of 30% in the NSB is negligible.

Note these results should be considered as a worse case scenario, as some 1.3*NSB gammas are lost in the gamma hadron separation, trained with 1.0*NSB

3. SENSITIVITY STUDIES FOR THE CTA

gammas and protons. Also, it should be taken into consideration that these were the first results of the NSB effect on high level performance (after data cuts) ever presented in the CTA collaboration, and proofed to be valid after the large-scale NSBx3 configuration was generated and analyzed.

3.6.4 La Palma alternative layouts

Taking into account the restrictions imposed by the terrain orography of the candidate site Roque de los Muchachos Observatory (ORM), alternative telescope positions with minor changes to the standard CTA North layout (“2N”, shown in Fig. 3.14) have been proposed in order to avoid some of the steepest areas. In this section two different layouts are tested, evaluating the effect in sensitivity and angular resolution with respect to the standard “2N” layout.

Proposed layout 2LP1 (shown in Fig. 3.35a) requires a maximum altitude difference of $\sim 50m$ and 2LP2 (shown in Fig. 3.35b) also requires convenient locations for construction, taking into account accessibility, slope and cliffs. These layouts are also restricted by the Prod-2 NORTH configuration limited number of simulated telescope positions, so these results should be considered just as lower limits.

Both Prod-2 Tenerife and Arizona (US) were analyzed using data with 20° of zenith angle, and comparing average performance for North and Southward pointing (same exposure time in each direction). In order to use identical training samples, the same subset of MC files was used for training the various reconstruction algorithms (e.g. energy reconstruction or background suppression). Under this considerations, the standard CTA-N layout “2N” was analyzed and compared with proposed “2LP1” and “2LP2” layouts. Note all these layouts have the same number of telescopes (4 LSTs and 15 MSTs), with 4 different MST positions each.

Comparing differential sensitivity (shown in Fig. 3.36) at the Tenerife site of analyzed layouts, no significant differences are observed between proposed and standard layouts with variations at the few percent level. Note that error bars only show statistical fluctuations, and may not be realistic as all layouts use the same MC data.

Comparing angular resolution (Fig. 3.37), differences appear above 10 TeV,

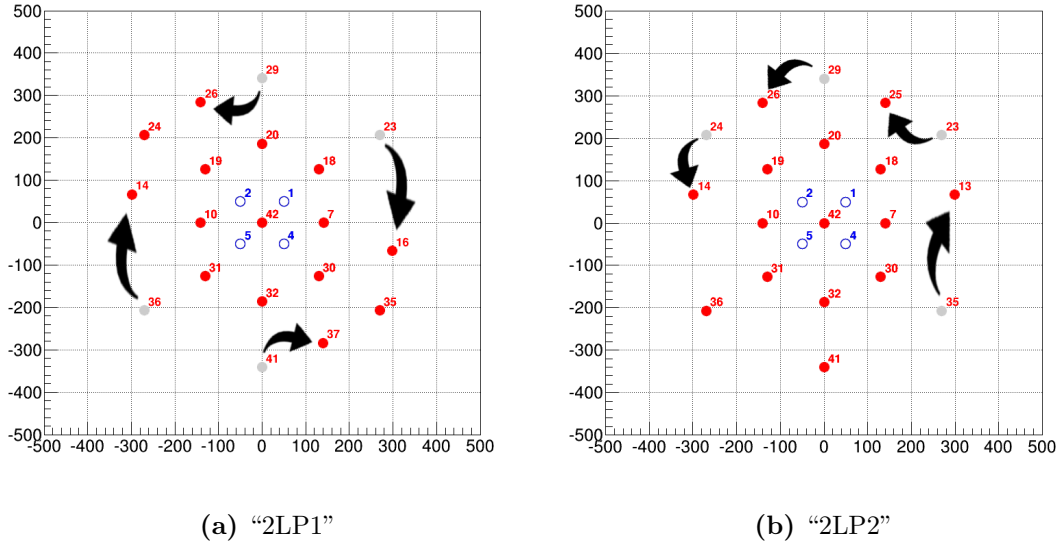


Figure 3.35: Proposed “2LP1” (*left*) and “2LP2” (*right*) layout candidates by A. Moralejo and J. Cortina. Empty circles correspond to LST positions while filled circles correspond to MSTs. Grey circles show standard “2N” layout positions and black arrows point to the new positions.

showing 2N layout outperforms proposed layouts, probably due to the MST symmetry. Within alternative layouts, 2LP2 seems to be the favoured option, with just a $\sim 15\%$ loss in angular resolution.

It must be taken into account that angular resolution differences have no effect on the differential sensitivity of a point-like source, which is reasonable, as at the highest energies CTA sensitivity is signal-limited, where wider cuts add no significant background. Note angular resolution makes use of the image axes, but does not use any image shape parameter or pixel timing, leading to a strong dependence between resolution and the observed stereoscopic angle. A more sophisticated reconstruction incorporating that additional information (like the one currently used in the MAGIC experiment) would improve the reconstruction of showers well outside the array, for which images are more or less parallel. It is reasonable to expect that the angular resolution differences become smaller with such an analysis.

As a cross-check, an identical analysis was performed using Prod-2 NORTH con-

3. SENSITIVITY STUDIES FOR THE CTA

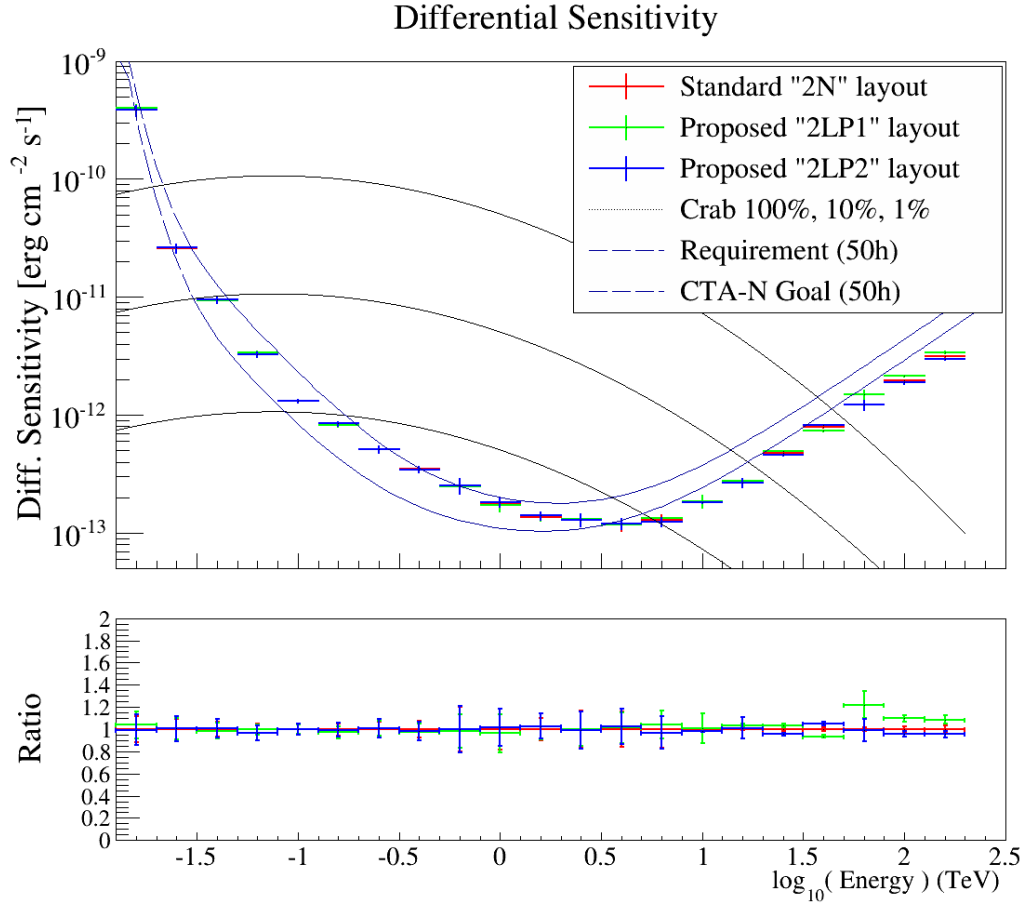


Figure 3.36: Differential sensitivity of the standard CTA-N “2N” and the two proposed alternative layouts. There is no significant loss in the sensitivity breaking the symmetry of the standard telescope layout.

figuration “USA” site data and similar results were obtained, with no significant differences in terms of differential sensitivity.

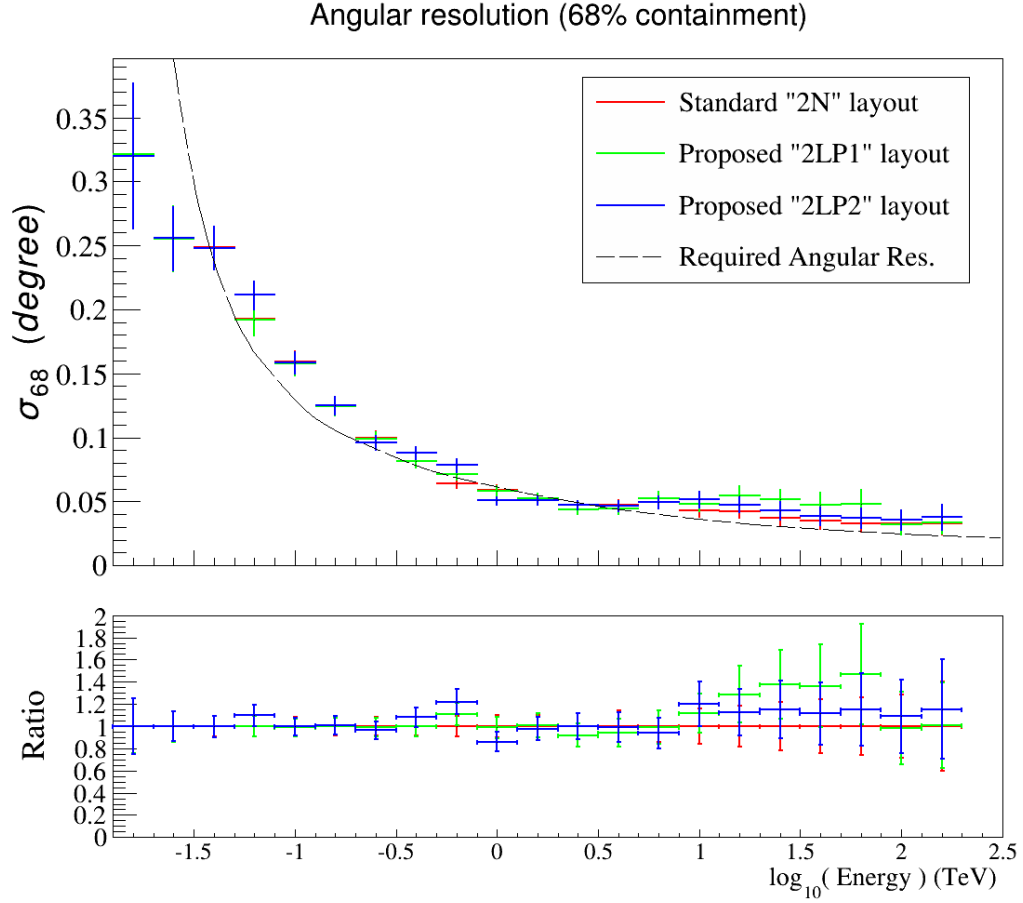


Figure 3.37: Angular resolution (corresponding to the 68% containment radius) of standard CTA-N “2N” and the two proposed alternative layouts. Candidate layout “2LP1” seems to be disfavoured above 10 TeV, showing a loss in angular resolution up to 40%.

3. SENSITIVITY STUDIES FOR THE CTA

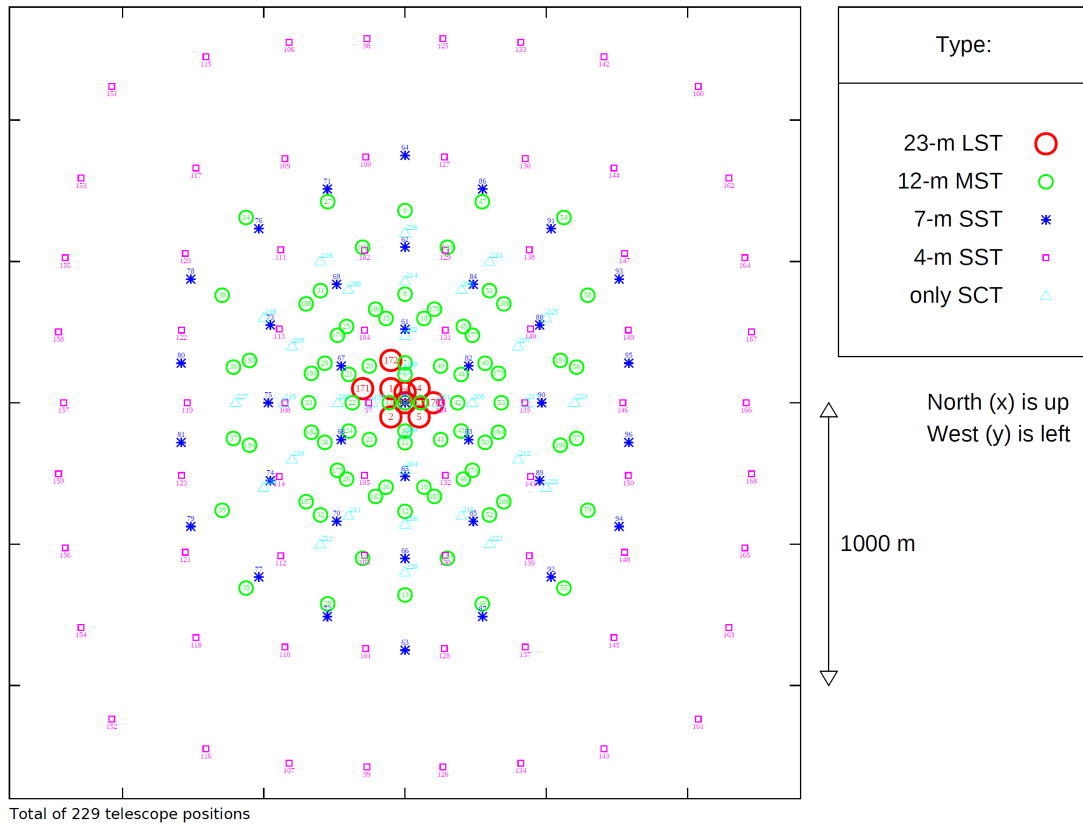


Figure 3.38: Standard Southern Hemisphere (of code-name STD) layout of simulated telescope positions (and types) of the second CTA MC production (Prod-2). Image courtesy of K. Bernlöhr.

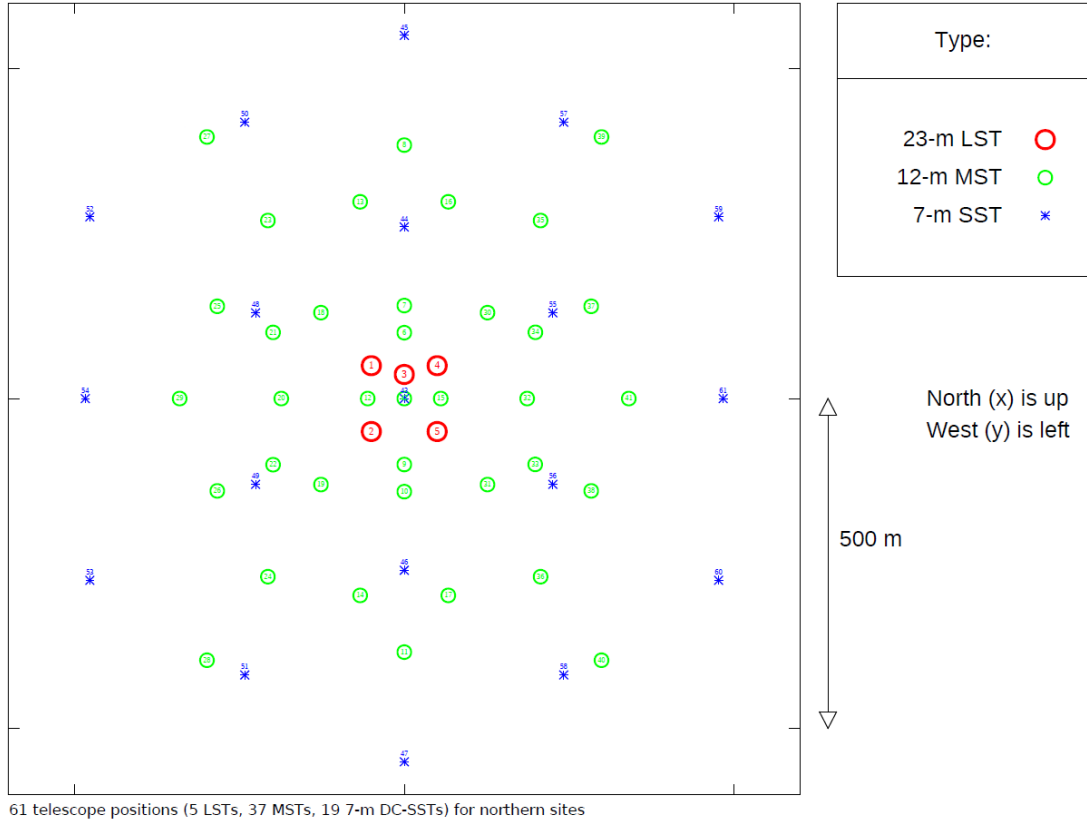


Figure 3.39: Standard Northern Hemisphere (of code-name **NORTH**) layout of simulated telescope positions (and types) of the second CTA MC production (Prod-2). Image courtesy of K. Bernlöhr.

3. SENSITIVITY STUDIES FOR THE CTA

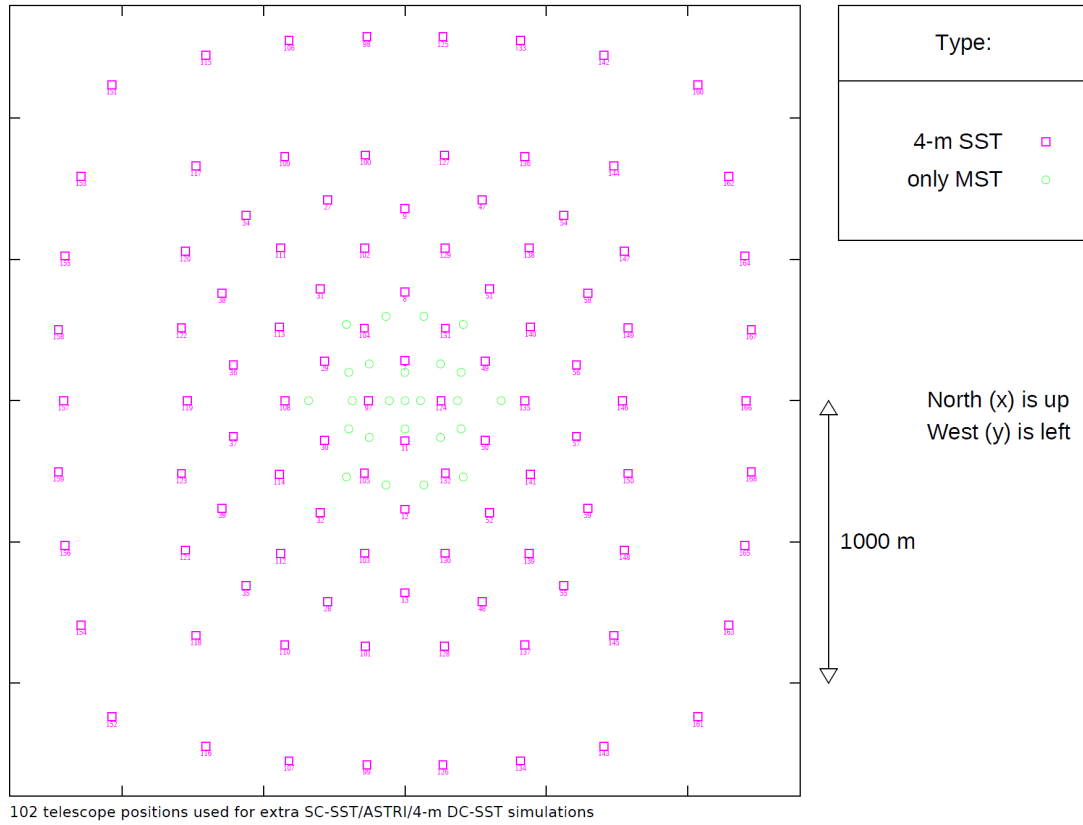


Figure 3.40: Simulated telescope positions for the 4m SSTs extension, equal for the 4MSST, SCSST and ASTRI configurations. Image courtesy of K. Bernlöhr.

Chapter 4

CTA forecast

Along with the evaluation of the CTA performance carried out through detailed MC analysis of described in the previous chapter, we aim to study its physics potential. As the observatory will possess capabilities far exceeding the current generation of IACTs, it would be impractical to use the population of sources observed from ground so far. It seems reasonable to take advantage of the extensive catalog of sources detected in the 30 MeV to 300 GeV range by *Fermi*-LAT. The final decision on a particular array layout or construction site can profit from these catalogs to understand the effect of the IRFs over the observatory potential scientific output.

In this section, populations of γ -ray sources extracted from the different catalogs published by the *Fermi* collaboration are used to assess CTA capabilities over on specific physics cases. Source spectra from each catalog are selected and extrapolated to the VHE regime using conservative assumptions. Predictions are then performed to assess the total observable population by the different CTA layouts.

With this goal in mind the author developed tools to estimate the detectability of simulated spectra by the CTA. Section 4.1 focuses on the different software packages used, comparing them to other alternative tools developed by the CTA collaboration. In section 4.2, the following physics cases are studied: *a)* the accessible galactic and extragalactic populations under different approaches, such as surveys or follow-up observations. *b)* The possibility of detecting other pul-

sars with similar VHE spectra as the one observed in the Crab pulsar [27]. *c)* Prospects for the detection of DM in dSph galaxies, providing estimations of the limits reachable on the WIMP particle cross section through different annihilation channels.

4.1 Physics evaluation tools

During the design phase of the CTA there was a need for an interface between MC simulations and the development of physics cases. Different groups developed tools to estimate CTA capabilities over specific test cases using MC generated Instrument Response Functions (IRFs), creating high level products such as the measured spectrum, light curves and sky-maps.

They are collected in the following list:

- *CTA macros*: After the definition of the γ -ray source spectrum, these tools estimate the significance level, SED, sky-maps or the comparison between different spectra performing a χ^2 -test. They may be the most complete and more cross-checked tools available, and were mainly developed by D. Mazin. They are written in C++ using the ROOT framework.
- *ctools*: Written in C++, they make use of the widely used Flexible Image Transport System (FITS) data format, and use a similar approach as NASA FTOOLS [150]. This package performs event selection and binning and performs model fitting through likelihood analysis. Generating spectra, SEDs and sky-maps is also possible. These tools may be the preferred option for people with experience with *Fermi*-LAT data analysis. They were mainly developed by J. Knödlseder and were licensed with a GNU GPL v3.
- *iCTA*: Creates ON and OFF event lists and produces source sky-maps, spectra and SEDs, storing them as FITS files, readable by classical astronomy software such as fv, DS9 or Xspec. Written in Interactive Data Language (IDL). *iCTA* is a private software mainly developed by M. Renaud.

All these tools were used by the collaboration to assess the impact on specific scientific cases of different candidate arrays simulated by the MC working group.

They were also employed in the definition of the Key Science Project (KSP)s. Simpler tools were developed by the author of this work to perform fast and accurate observability prospects. They are described in the next section.

4.1.1 *GAETools*

A set of tools were created by the Grupo de Altas Energías (GAE) group to perform a fast significance level calculation from a simulated γ -ray spectrum emitted by a point-like source. This software was initiated by R. García, N. Mirabal, E. Currs and J.L. Contreras and further developed by the author of this work.

To calculate the attained significance, N_{on} (number of events in the *On* region) and N_{off} (number of events in the *Off* region) are estimated, using the IRFs, generated in the MC analysis. Specifically, the effective area and background rates (as a function of the energy) are used. The calculation of the significance of the detection of a point-like source with a spectrum $F(E)$ at redshift z , within a certain energy range ($E_0 < E_{\text{obs}} < E_f$), observed during t hours goes as follows:

- Total background (BG_T): The background rate is extracted from the IRFs and integrated between E_0 and E_f . This rate is then multiplied by the observation time t , obtaining the total number of background events observed by the telescope in the *On* region.
- EBL absorption: Making use of the redshift of the source, z , and one absorption model (described in Sec. 1.2.1) the transmission as a function of the energy $\tau(E)$ is computed. By default, Franceschini model [12] is used.
- Excess rate ($N_{\text{rate}}(E)$): The excess rate as a function of the energy is calculated multiplying $F(E)$ by the effective area and the transmission factor $e^{-\tau}$:

$$N_{\text{rate}}(E_i) = F(E_i) * A_{\text{eff}}(E_i) * \tau(E_i)$$

- Total excess (N_T): Integrating the excess rate between E_0 and E_f and multiplying it by the observation time t , the total number of events observed from the source within the observation time and the considered energy range

4. CTA FORECAST

is calculated. The integration is performed in discrete steps using the *trapezoidal rule*:

$$N_T = t \cdot \sum_{E_0}^{E_f} (E_{i+1} - E_i) \frac{N_{rate}(E_i) + N_{rate}(E_{i+1})}{2}$$

- Significance (S): To estimate the significance, Eq. 3.2 (equation 17 from [136]) is used with the expected number of events in the *On* and *Off* regions calculated from the previous steps. Within the *On* region, signal from the source and background are expected; $N_{on} = N_T + BG_T$. In the *Off* region only background is expected and the ratio between the observation times dedicated to the *On* and *Off* regions needs to be considered, making use of α ; $N_{off} = \alpha \cdot BG_T$. The α value generally used corresponds to a number of background-control regions of 5, that is $\alpha = 0.2$ in the Li&Ma notation. This value may be considered conservative, given the wide FoV of the CTA, which may allow for even lower values (more background-control regions).

Additional requirements to the standard condition of $S > 5\sigma$ are imposed before assuming a detection. The total excess N_T is required to be larger than the total background times the level of possible systematic effects considered (a conservative 5% in our case): $N_T > BG_T * 0.05$. N_T is also required to exceed 10 events. If all conditions are accomplished, the source is classified as detectable by the selected candidate array within the observation time used.

The observation time generally set as default is 50 hours in order to be as close as possible to the time considered in the IRFs calculation within the MC analysis. From the estimation of the observed significance of a source in 50 hours, the observation time required for a detection (t which makes $S = 5\sigma$) can be inferred, taking into account that for a perfectly known background the significance grows proportionally with the square root of time:

$$t_{5\sigma} \simeq 50 \frac{5^2}{S_{50h}^2} \text{ hours} \quad (4.1)$$

where S_{50h} is the calculated significance in 50h of observation and $t_{5\sigma}$ the minimum observation time required for a detection. Note this should be treated as an

approximation, as large differences between the calculated $t_{5\sigma}$ and the observation time used in the MC analysis (usually 50 hours) would lead to inconsistencies. These inconsistencies come mostly from the fact that this relation is too simplistic for the CTA low energy range, where the minimum observable flux may be limited by the level of systematics considered.

The software needed to be flexible in order to produce accurate prospects for different physics cases. Regarding pulsars, a few more considerations were taken into account, as the analysis allows for an additional background reduction. Pulsars emit γ -rays periodically during small time phases, around the light curve peaks while background events are uniformly distributed over time. These intervals are selected, dumping events which are not coherent with the pulsed emission. The ratio between the selected time windows and the total phase time is generally called *duty cycle*. Typical values of $\approx 10\%$ were assumed, corresponding to an average background rejection of 90%.

4.1.2 Comparison with *CTA macros*

Several comparisons were performed between *CTA macros* and *GAETools*. It was not possible to crosscheck the package against other ones because they were not available at the time or they were not public. Equal spectra were generated and the significance was calculated using both methods, taking care of integrating the signal over similar energy ranges, using the same CTA IRFs. Results were totally correlated at low significance levels (between approximately 3 to 10 σ) and would differ slightly at higher values (above $\approx 20 \sigma$). As the main objective of the tools is to perform detectability prospects, deviations at high significance would have no effect on the conclusions, only affecting slightly the calculation of the observation time needed to reach a 5 σ detection.

Fig. 4.1 shows a comparison of the sensitivities obtained by *CTA macros* and *GAETools* over the same sample of spectra, corresponding to the AGNs present in the 1FHL catalog, extrapolated to the VHE range. A good correlation is found between both methods.

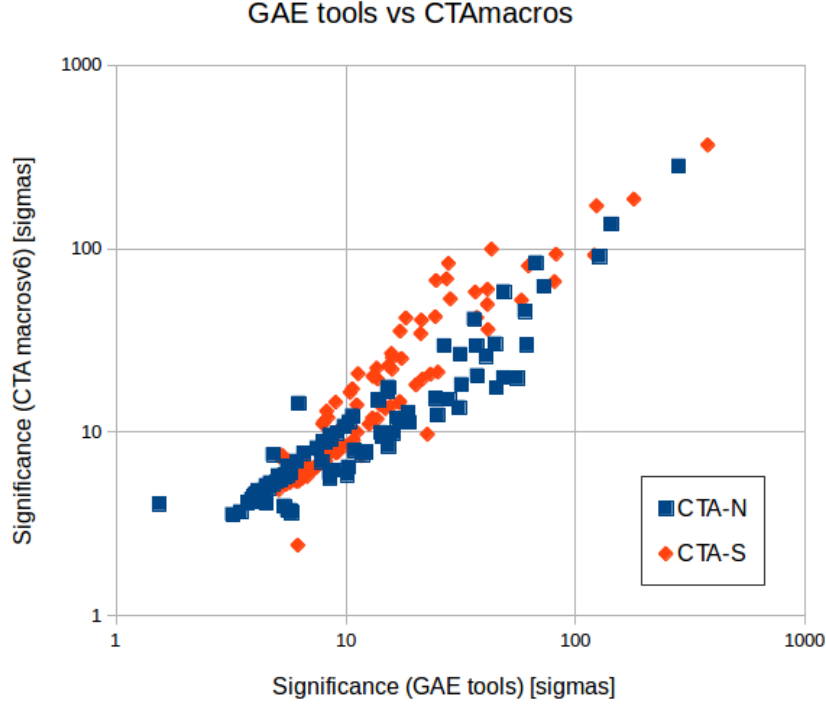


Figure 4.1: Comparison of calculated significances between *CTA macros* and *GAEtools*. The source sample corresponds to the AGNs present in the First *Fermi* LAT Catalogue of Sources Above 10 GeV (1FHL) observable by CTA-N or CTA-S.

4.1.3 *Fermi* catalog extrapolations

GAEtools were developed to build up prospects of the expected source populations detectable by CTA. In this work, populations of γ -ray sources extracted from the different catalogs published by the *Fermi* collaboration were used to assess CTA performance on specific physics cases and compare different candidate arrays. With this purpose in mind, extrapolations to higher energies were performed using the 1451 sources present in the First *Fermi* LAT Catalogue (1FGL) and later updated to the 1873 objects within the 2FGL. The recent 3FGL [24] has just been published during the time of writing so the results shown in this chapter could be updated in the future, although the main conclusions obtained are not expected to change with larger populations.

The 1FGL catalog modeled all sources with a simple power-law, although there

were clear discrepancies with data in the case of bright sources. With the increased exposure accumulated discrepancies grew larger in the 2FGL, making necessary to use different models. A big fraction of both galactic and extragalactic sources can be well described by a power-law:

$$\frac{dN}{dE} = K \left(\frac{E}{E_0} \right)^\Gamma \quad (4.2)$$

where K is the flux at E_0 and Γ is the power-law index (slope). Most of the remaining sources were modeled using a log-parabola:

$$\frac{dN}{dE} = K \left(\frac{E}{E_0} \right)^{-\alpha - \beta \log(E/E_0)} \quad (4.3)$$

where α is the spectral slope at E_0 and β the curvature index. Pulsars within the 2FGL are modeled with an exponential cutoff power-law:

$$\frac{dN}{dE} = K \left(\frac{E}{E_0} \right)^\Gamma \exp \left(-\frac{E - E_0}{E_c} \right) \quad (4.4)$$

where E_c is the *cutoff energy*.

A few extra considerations were introduced in the *GAETools* analysis:

- Inverse Compton peak: In agreement with the observations measured by the current generation of γ -ray detectors, hard spectra (i. e. spectral shapes with relatively flat slopes and high VHE emission) are expected to soften at higher energies above the inverse Compton (IC) peak. Hard spectra measured by *Fermi*-LAT fall before the IC peak, and the extrapolation to higher energies needs to be corrected. Hard power-law spectra ($\Gamma < 2$) are artificially softened using a broken power-law, introducing a spectral break at 100 GeV (shifted with the redshift a factor $1/(1+z)$) changing their slope to $\Gamma = 2.5$. Spectra modeled with log-parabolas or with exponential cutoffs already soften at higher energies, so no correction is needed.
- Pulsars: As described in 4.2.3, the exponential cutoff describing pulsars in the 2FGL is contradicted by recent observations of IACTs [25, 26, 27], extending the Crab pulsed emission up to 400 GeV following a power-law. The

4. CTA FORECAST

possibility of a similar VHE component being present in other pulsars is studied in this work, using broken power laws of variable slope as pictured in Fig. 4.7.

Specific observability forecasts were performed using the AGN sample present in the First *Fermi* LAT Catalogue of Sources Above 10 GeV (1FHL) [151] helping to define the strategy of observation of blazars within the AGN KSP. The approach described of using broken power-laws above 100 GeV was employed. It must be noted that these spectra are modeled using *Fermi*-LAT data points from 10 GeV up to 300 GeV, and already suffer from some EBL absorption, producing a softening of the observed spectra for far away blazars. Performing extrapolations from these spectra without deabsorption should be considered a conservative approach, as spectra will be artificially softened and EBL absorption will be applied to calculate significance.

4.1.4 DAMASCO

Dark Matter (DM) prospects were also performed, integrating within *GAETools* the DArk Matter Analytical Spectral COde (DAMASCO) package, a ROOT-based implementation of dark matter WIMP annihilation predictions for photon spectra for the most frequently considered dark matter annihilation channels [152]. This package was mainly developed by D. Nieto and the author of this work.

Spectra are generated by DAMASCO (see Fig. 4.2), and several calculations are performed by the *GAETools*:

- **Detectability:** An evaluation of the statistical significance of the DM signal is performed as a function of the WIMP particle mass m_χ and the astrophysical factor (J) for different possible annihilation channels. The minimum astrophysical factor J_{\min} required to reach a statistical significance of 5σ is calculated, assuming a certain effective observation time, and the thermal cross-section $3 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$. As the expected spectrum of WIMP annihilation channels are considered known, the significance is maximized by selecting the low energy cut which achieves the best statistical significance. Upper limits should be treated as a first approximation, as more detailed

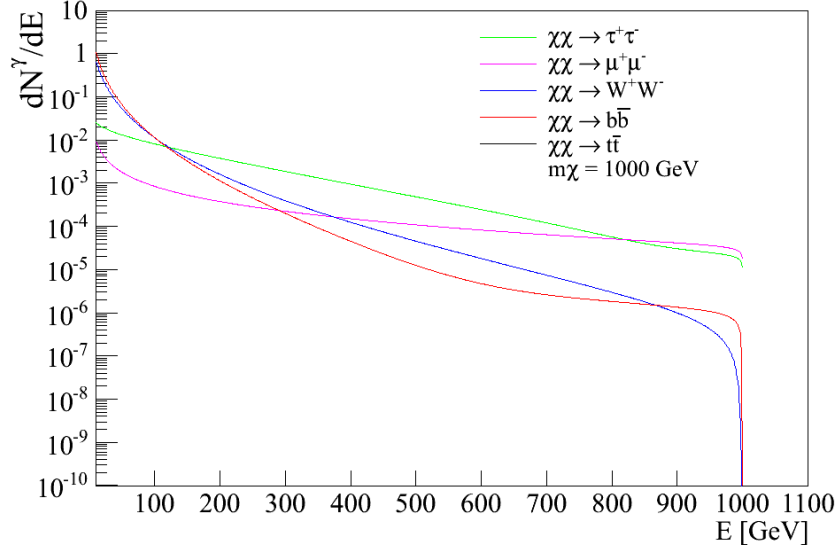


Figure 4.2: γ -ray spectra from the annihilation of 1 TeV WIMPs for the $\tau^+\tau^-$, $\mu^+\mu^-$, $b\bar{b}$, and W^+W^- channels from [152], generated with the DAMASCO software package.

knowledge on the signal and background distributions (as in [153]) could improve results.

- **Annihilation cross-section:** Bounds on the annihilation cross-section of WIMP particles are calculated from the upper limits attained by CTA, inferred from the methods described in [154], with a confidence level of 95%. The sensitivity is calculated assuming that the DM particle annihilates purely into each channel, a certain observation time and a solid angle $\Delta\Omega$.

Results obtained by using this software are shown in section 4.2.4. More specific details on dSph galaxy searches using IACTs can be found in [10].

4.2 Physics cases

This section collects the results obtained applying the IRFs deduced from the simulation of CTA layouts to specific physics cases.

With 24 months of accumulated data, the 2FGL catalog contains 1873 point sources characterized in the 100 MeV to 100 GeV energy range. Many of these

4. CTA FORECAST

sources have not yet been discovered in the VHE regime probably due to the lack of sensitivity of existing instruments, but they are expected to be accessible to the CTA. As an example, about 85% of the VHE active galactic nuclei detected by the current ground-based Cherenkov experiments are found in the 2FGL. Extrapolating sources present in this catalog to higher energies seems to be a sensible procedure for the compilation of a tentative catalog of future CTA detectable sources.

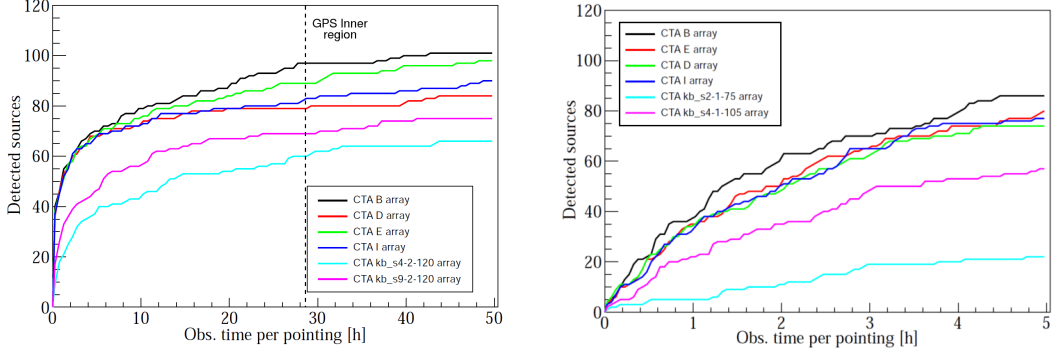
Using *GAETools*, the software described in Sec. 4.1, several physics cases were studied in order to assess the CTA capabilities, compare Prod-1 candidate arrays (see Table 3.3 for reference) and Prod-2 building sites (see Table 3.5) performance. For simplicity, full CTA layouts were assumed to have full sky coverage, making no distinction between north and south (with the exception of the observability study performed in Sec. 4.2.2.1). The results presented could improve if more recent IRFs were used. Nonetheless, they already show the great capabilities of the observatory, giving a general overview of the CTA scientific potential.

The studied physics cases are the following: *galactic* and *extragalactic* populations accessible to the CTA, pulsars detectability prospects, and the observation of dSph galaxies exploring the possibility of indirect Dark Matter (DM) detection.

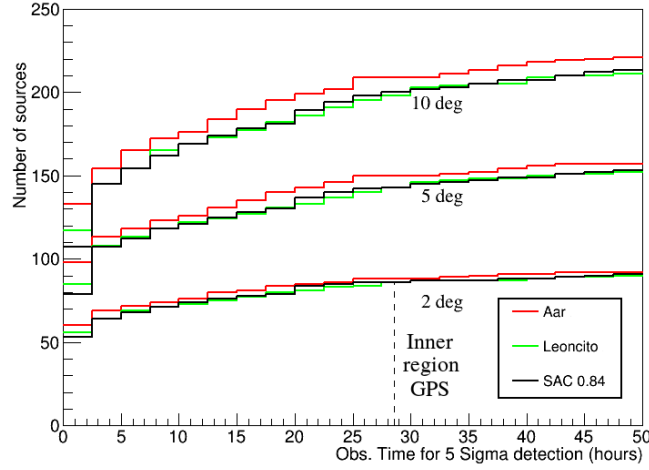
4.2.1 Galactic sources and surveys

To assess CTA capabilities for galactic source detection and the potential of galactic and extragalactic surveys, the reported 2FGL spectral parameters for the 1873 sources it contains were extrapolated to the VHE range (15 GeV – 300 TeV) [155]. As described in section 4.1.3, for each individual source the corresponding power-law or log-parabola parameters prescribed in [156] are adopted. Once the extrapolated spectra are fixed, detectability was tested using the *GAETools*, described in Sec. 4.1.1, using the CTA IRFs for different telescope configurations and site locations. For simplicity, 5 off-regions for each on-region observation, a 5% systematic error and standard detectability conditions were adopted.

The sample of galactic sources was constructed from all associated/unassociated sources with galactic latitude $b < 2$. It includes high-mass binaries, SNRs, Pulsar Wind Nebulae (PWN) and unassociated sources. In a second step, sources labeled



(a) Galactic sources ($|b| < 2^\circ$) (Prod-1 layouts) (b) Extragalactic sources (Prod-1 layouts)



(c) Galactic sources ($|b| < 2^\circ, 5^\circ$ & 10°) (Prod-2 layouts)

Figure 4.3: Cumulative distributions of the number of *Fermi*-LAT galactic and extragalactic sources detectable by the CTA at the 5σ level, as a function of individual pointing observation time. The equivalent time of observation to reach in the inner region of the CTA Galactic Plane Survey (GPS) KSP (28.6 h) is highlighted. Several CTA configurations were tested: *Top Left:* Galactic ($|b| < 2^\circ$) source detectability by Prod-1 candidate layouts (arrays B, E, D, I, “kb_s4-2-120” and “kb_s9-2-120”). *Top Right:* Extragalactic source detectability by Prod-1 candidate layouts (arrays B, E, D, I, “kb_s2-1-75” and “kb_s4-1-105”). *Bottom:* Galactic detected populations for different explored widths of the galactic plane (2° , 5° and 10°) by Prod-2 “2A” layout simulated at several site locations of different altitudes: “Aar” (1640 m), “Leoncito” (2662 m) and “SAC” (3600 m). Note “kb_s2-1-75”, “kb_s4-1-105”, “kb_s4-2-120” and “kb_s9-2-120” layouts correspond to subsets of 2 LSTs, 4 LSTs, 4 MSTs and 9 MSTs respectively.

4. CTA FORECAST

as extragalactic together with highly variable unidentified sources were excluded. This strategy resulted in a total of 196 tentatively tagged sources in the Galactic category at $|b| < 2^\circ$. Repeating this exercise for sources at $|b| < 5^\circ$ increases the initial sample to 297, and imposing $|b| < 10^\circ$ to 390. Given the limited spatial information from the 2FGL, the entire sample was treated as point-like sources. As described in section 4.2.2, an independent sample is created with extragalactic sources in order to perform analog studies.

The key variable to consider is the equivalent observation time employed per on-axis pointing while performing a sky survey. The resulting observable population as a function of the individual pointing observation time is shown in Figure 4.3a. It shows that ≥ 70 2FGL galactic sources (or 35% of the initial samples) would be detected by CTA with individual exposure times of 5 hr or more when using full-array configurations (B, D, E or I). The performance of smaller subsets of the array (“s4-2-120” and “s9-2-120” with 4 and 9 medium-sized telescopes, respectively) was also considered. While not as effective as a fully dedicated array, the fraction of detected sources remains significant. One option would be to use large-sized telescopes for extragalactic sources (which tend to have softer spectra, due to the EBL absorption) and small/medium-sized telescopes for the Galactic Plane sources (which tend to have harder spectra).

Although this is not possible within a survey strategy, note that up to $\approx 50\%$ of the 2FGL sample of Galactic sources are within the reach of CTA, using exposure times as long as 50 h/source (Tab. 4.1). The best array configurations for this are B and E, due to their increased number of LSTs. With the observed *Fermi*-LAT source density in the Galactic plane ($|b| < 5^\circ$) and a 25 square degree FoV gives an average of 2 sources per field, a total of 4000 hr would therefore be needed to complete a targeted survey returning 50% of the *Fermi*-LAT Galactic catalog. Regarding the altitude of constructed site, as depicted in 4.3c, moderate altitudes seem favoured as the Namibian site (1640 m) outperforms alternative candidates with an improved coverage between 5 to 10%.

The possibility of performing an extragalactic survey was also explored, using the expected effective areas and background rates from CTA Prod-1 candidate arrays. Results show that a CTA all-sky survey with typical exposure times of 0.5 h would detect only ≈ 20 *Fermi*-LAT extragalactic sources over the whole

Array	Extragalactic		Galactic ($ b < 2^\circ$)	
	5 h	50 h	28.6 h	50 h
B	89	192	97	101
D	74	138	79	84
E	80	171	89	98
I	77	159	83	90
2 LSTs	21	72	-	-
4 LSTs	56	135	-	-
4 MSTs	-	-	60	66
9 MSTs	-	-	69	75

Table 4.1: Number of *Fermi*-LAT sources selected from the 2FGL catalog that would be detected by CTA using different individual on-axis exposure times up to 50 h and various Prod-1 array configurations (see Figs. 4.3a-4.4a). *Extragalactic*: sources accessible to CTA in 50 h, and those CTA would detect with an all-sky extragalactic survey with an equivalent exposure time of 5 h. *Galactic*: standard detectable sources within 50 h, and those detectable with a survey through the galactic plane (with $|b| < 5^\circ$) with an equivalent exposure time of 28.6 h, as proposed for the inner region of the GPS KSP.

sky (see Fig. 4.3b), *i.e.* 5 within the 1/4th of the sky observable with good zenith angle (assuming the sources are distributed uniformly over the sky). With 5 h/pointing the number of sources increases to 80 (*i.e.* 20 in practice). The total number of *Fermi*-LAT extragalactic sources detectable with CTA exposure times up to 50 h/source is ≥ 170 (30% of the initial sample) with the most favorable array configurations (B and E, Tab. 4.1 and Fig. 4.4a).

4.2.2 Extragalactic follow-up observations

In addition to an extragalactic survey, performing follow-up observations of known sources already detected at lower frequencies offer another promising approach in the search of AGNs in the VHE regime. The detected blazars in the *Fermi*-LAT energy range, which overlaps with the low energy threshold of CTA, provide valuable information of the intrinsic spectral properties of the extragalactic sources that will be accessible to the observatory. Future AGN concurrent observations by *Fermi*-LAT, CTA and HAWC will provide a broad-band measurement over more than 5 decades in energy.

To produce a conservative estimate, 2FGL AGNs were selected as long as they

4. CTA FORECAST

Table 4.2: Number of detectable *Fermi* extragalactic sources with known redshift for different array configurations (50 h of exposure time). AGNs with unknown type are classified as “other AGN”.

Array	FSRQs	BL Lacs	other AGN	SBGs	RGs	Seyferts	Total
Prod-1							
B	46	117	19	3	6	1	192
C	17	84	17	3	6	1	128
E	32	111	18	3	6	1	171
NA	33	109	18	3	6	1	170
NB	27	103	17	3	6	1	157
Prod-2							
2A (Namibia)	53	130	20	3	6	1	213
2A (Leoncito)	50	122	19	3	6	1	201
2A (SAC 0.84)	51	124	19	3	6	1	204

were labeled as extragalactic objects and they had confirmed redshifts in the Bz-cat or Véron (13th edition) catalog. 561 sources from the cross-correlated 2FGL catalog were selected for further processing, including 340 FSRQs, 171 BL Lacs, 10 radio galaxy (RG)s, 6 Seyfert galaxies, 3 starburst galaxies (SBGs) and 31 AGNs of other classes.

In order to make predictions for CTA, similarly as in the previous section, the final subset of 561 *Fermi* sources were extrapolated to higher energies as described in Sec. 4.1.3. The EBL absorption as a function of their redshift [12] and the expected softening at higher energies of very hard sources ($\Gamma < 2$) were taken into account. Attainable significance for different observation times was calculated using *GAEmacros*, with an α value of 0.2 and requiring a total excess over 5% of the background. To maximize the number of detections in this study and assess CTA capabilities more accurately, several energy thresholds were used to optimize the signal and significance estimation of each source. This method would be analog to the low/medium/high energy cuts optimization performed in the current IACT experiments.

The performance of different array configurations was explored using CTA Prod-1 candidate arrays. Different altitudes were also tested by comparing Prod-2

proposed layout “2A” simulated at different construction sites: “Aar” (1640 m), “Leoncito” (2662 m) and “SAC” (3600 m). As a result, as shown in Table 4.2, more than ≥ 150 extragalactic objects would be typically detected in 50 hours of maximum exposure time per source. When Prod-2 “2A” layout was considered, those numbers augmented to ≥ 200 . Note that CTA will be most efficient for hard sources ($\Gamma < 2$) and shall reveal all the complexities of the extragalactic population extending beyond the reach of current detections. For softer sources that currently are only accessible during flares, CTA will provide unique access to quiescent states, and measure with improved sensitivity and time resolution their flaring states.

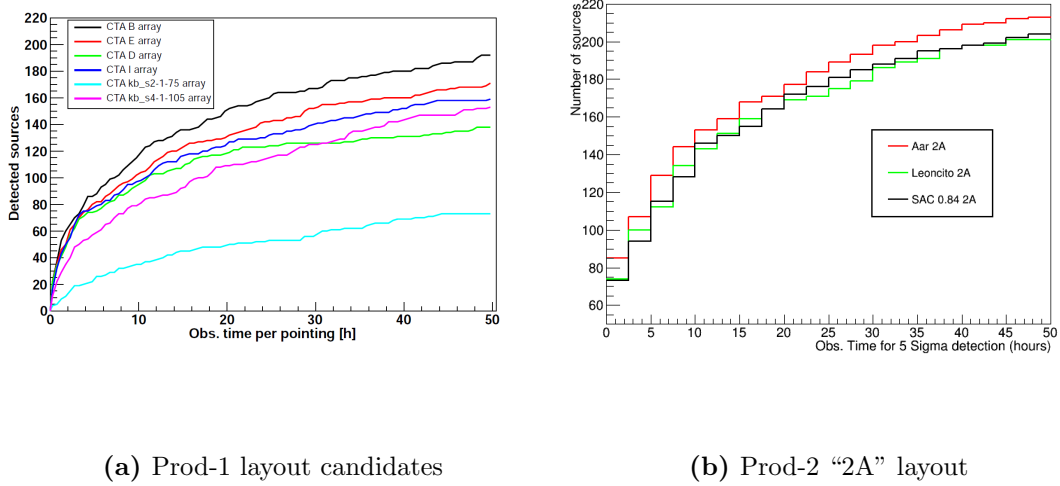


Figure 4.4: Cumulative distribution of the number of *Fermi*-LAT extragalactic sources of known redshift detectable by the CTA at the 5σ level, as a function of the individual pointing observation time. Several CTA configurations were tested: *Left*: Some of the Prod-1 candidate layouts (arrays B, E, D, I, “kb_s2-1-75” and “kb_s4-1-105”). *Right*: Prod-2 “2A” layout simulated at several site locations of different altitudes: “Aar” (1640 m), “Leoncito” (2662 m) and “SAC” (3600 m). Note “kb_s2-1-75” and “kb_s4-1-105” layouts correspond to a subset of 2 LSTs and 4 LSTs respectively.

The Prod-1 array configurations that do not include LSTs, such as array “C”, yield the poorest performance. They give access to significantly fewer FSRQs, which usually have very steep spectra, but the same is also true for BL Lacs. Ar-

4. CTA FORECAST

ray “B”, with the best coverage at low energies, yields the best results in terms of source statistics. Note these results exclusively measure detectability, and do not gauge the spectral energy range in which these sources are significant. The “compromise” solutions, such as configuration “E” and the northern array “NA”, remain as excellent options with a slightly lower number of detections, but improved VHE sensitivity. Regarding the site altitude, the Namibian site at an altitude of 1640 m outperforms alternative candidates with a relative improvement of 5%. As shown in Fig. 4.4a and 4.4b, comparing the amount of sources detected at different individual observation times clearly demonstrates how layout performance improvements reduce the required observation times to reach an equivalent source population.

Apart from augmenting the total number of extragalactic TeV sources, CTA should also increase dramatically the number of objects visible at high redshifts. Fig. 4.5 shows the distributions of detectable AGNs as a function of the redshift. The most distant quiescent detectable AGN predicted are at $z \approx 2$, although this limit will likely be raised through observations during flares as most distant VHE AGNs detected by the current generation of IACTs were in these states. Note that sites with higher altitudes would allow an improved low energy threshold, unveiling more far away blazars.

These results illustrate the remarkable capabilities of CTA compared to current IACTs. The detection of at least 200 extragalactic sources is expected in less than two years with a maximum of 50 hours per source. Actually, the total number of *Fermi* AGNs detectable in 50 hours would reach about 400 if the 2FGL BL Lacs without known redshifts were detected with the same efficiency as the ones used in this study. In fact, the main difficulty in elucidating the full BL Lac population lies in obtaining direct redshift measurements from their mostly featureless optical/UV spectra. A possible workaround might come from a direct measurement of the shape of the EBL, which would allow us to set upper limits on unknown redshifts. The artificial break at 100 GeV for hard sources might also lead to an underestimation of the number of detections, but on the other hand, some of the sources might have intrinsic spectral breaks or cutoffs above the *Fermi*-LAT energy range, which would reduce their signal in the VHE band. Only CTA will be able to inform us on the actual spectra beyond 1 TeV for most of these sources. It

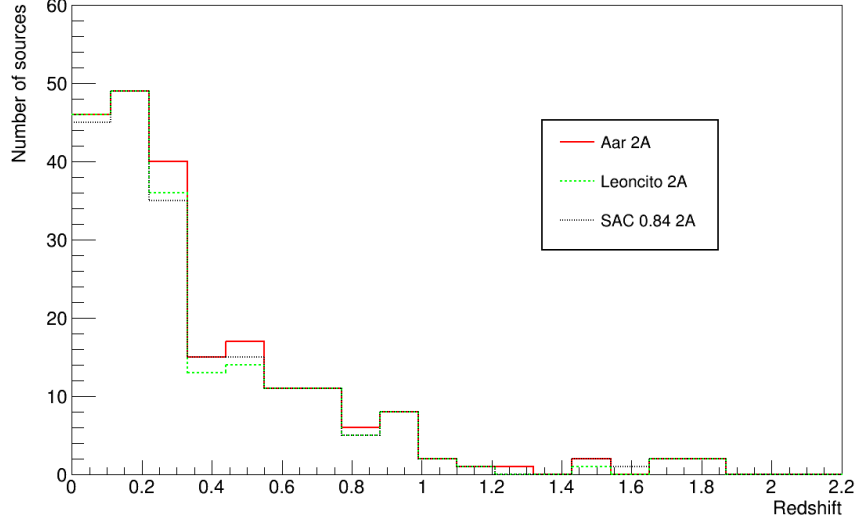


Figure 4.5: Number of *Fermi*-LAT extragalactic sources of known redshift detectable by the CTA at the 5σ level as a function of the redshift. Different curves correspond to the CTA-S candidate array “2A” from Prod-2 simulated at several site locations of different altitudes: “Aar” (1640 m), “Leoncito” (2662 m) and “SAC” (3600 m).

must be also noted that in addition, CTA would cover a wide area (roughly, 8% of the whole sky) with high chances of serendipity detections.

Discoveries such as extreme AGNs or extragalactic sources not adapted to the *Fermi*-LAT band and with emission peaking in the CTA energy range should further increase the sample of AGNs seen by CTA. Given that currently 6 out of the 45 TeV AGNs detected from ground were not yet present in the 2FGL, one can anticipate a fraction of about 15% of additional CTA sources when making predictions based on this catalog.

4.2.2.1 CTA-N impact

The CTA-N observatory will allow full sky coverage and will likely be devoted mostly to extragalactic science. In this section we test the impact of building CTA-N on the observable AGN population extrapolating the 1FHL catalog, composed only by sources detected by *Fermi*-LAT above 10 GeV. These sources are very likely to be observed by CTA, and the extrapolation considered should be

4. CTA FORECAST

treated as conservative, as the spectra are extrapolated from a SED which already suffered EBL absorption. Note that having a broad population of detected AGNs at different redshifts will not only help to unveil their nature and emission mechanisms but will also help to significantly constrain the EBL, obtaining information about structure formation processes throughout the history of the universe [15].

At the time this work was written, there was no final decision on the location of the future observatory, so approximate values to the candidate sites location were used: For CTA-N and CTA-S, latitudes of $+30^\circ$ and -25° were used respectively. Source observability is calculated taking into account their equatorial coordinates. All sources with declinations 35° away or closer to the CTA assumed site latitudes are considered. At the time of this work there were no IRFs for $\pm 40^\circ$ of zenith angle and far away AGNs are only detectable close to the low energy threshold. If higher zenith angles were considered as observable, results would be too optimistic.

IRFs from the average performance of $\pm 20^\circ$ in zenith angle of “2N” and “2A” candidate layouts simulated at Tenerife and Namibia sites were used as CTA-N and CTA-S respectively, as they were the chosen simulations to evaluate CTA KSPs by the consortium.

Table 4.3: Observable AGNs from the 1FHL that would be detected in 50h of observation by CTA-N and CTA-S.

Observable AGNs	Site			
	Any	CTA-N	CTA-S	Both
All	172	100	92	20
$z > 0.5$	52	31	30	9
$z > 1$	15	8	10	3

Table 4.3 shows the great impact on the observable (and detectable in 50 hours) AGN population obtained by adding CTA-N. Bear in mind that the performance of the northern hemisphere uses “2N” candidate layout with a significantly smaller array of telescopes, and still manages to increase by more than a 50% the AGN population. This effect may seem unrealistic, but it is partially due to the fact that the presence of the galactic plane covering a significant part of the southern hemisphere sky decreases the amount of detectable AGNs by γ -ray detectors.

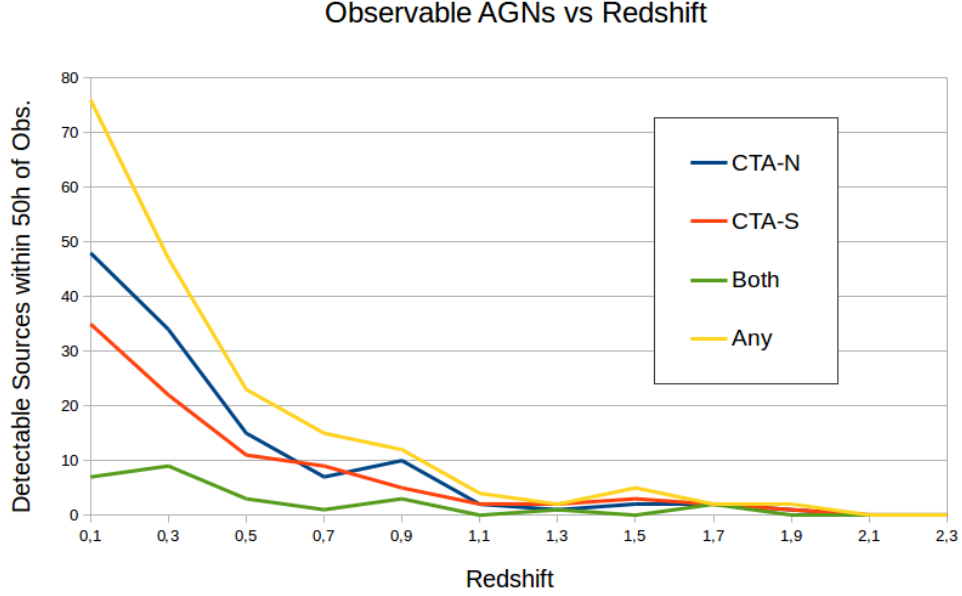


Figure 4.6: Observable AGNs from the 1FHL catalog that would be detected in 50 hours of observation by CTA-N and CTA-S as a function of the redshift. Both detectability and observability were considered, taking into account a flat performance over the zenith angle. Sources were considered observable if they peaked 35° away or closer to the zenith angle with respect to the north and south observatories at latitudes of $+30^\circ$ and -25° respectively. IRFs from Tenerife and Namibia candidate layouts “2N” and “2A” at $\pm 20^\circ$ of zenith angle were used respectively.

Fig. 4.6 shows the number of observable AGNs detectable within 50 hours of observation present in the 1FHL as a function of their redshift. The CTA-N and CTA-S detected sources are pictured separately, along with the number of sources detected by any of CTA sites and those accessible to both of them. The comparison between CTA-S and “Any” curves shows the great impact CTA-N will have on extragalactic astronomy, also increasing the amount of distant AGNs, observing as far as redshifts with $z \approx 2$ in their quiescent state.

4.2.3 Pulsars

The detection of neutron stars through gamma-ray pulsations is a key science goal for CTA. Gamma-ray pulsar observations at high energies (over a few tens of GeV) could help to understand the region where pulsed emission takes place by

4. CTA FORECAST

comparing the measured spectra with predictions by theoretical models.

The *Fermi* mission has revolutionized the study of gamma-ray pulsars detecting more than 117 sources in the MeV-GeV energy range [157], whose spectra are reasonably fitted with exponential cutoff values between 0.7 to 7.7 GeV.

Nonetheless, the detection of the Crab pulsar above 25 GeV with IACTs [25, 26] has reframed the exponential cutoff observed by *Fermi* in favor of a broken power-law shape that extends the pulsed emission up to 400 GeV. This recent discovery motivates the need for further pulsar studies in the VHE regime.

To place this situation in context, Fig. 4.7 shows the spectral fits (power-law with exponential cutoff) for 46 *Fermi* pulsars taken from the 1FGL, in comparison with the standard CTA differential sensitivity curve for the Prod-1 configuration “B” (5 LSTs) in 50h. The fits of Vela, Crab and Geminga pulsars are indicated explicitly, while the shaded area contains the fits of the remaining 43 pulsars.

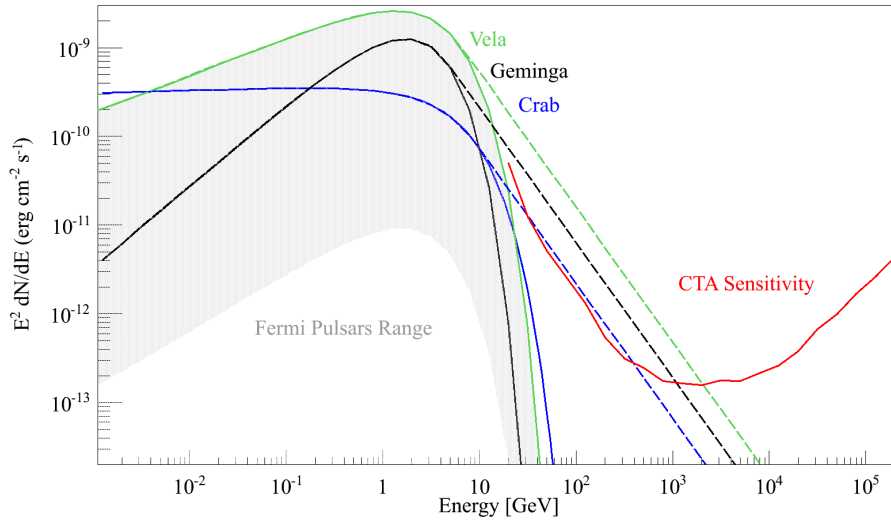


Figure 4.7: *Fermi*-LAT pulsars general profile (grey area) with Prod-1 CTA sensitivity curve for configuration “B” in 50h of observation. Vela, Geminga and Crab show extrapolated SEDs (dashed lines). Note the CTA sensitivity curve is shown for reference, as it does not account for the background reduction applied within pulsars analysis. Also note that a curve below the differential sensitivity curve can still be detected if the integral flux is high enough.

Initially, a 50 h simulated observation of the Crab pulsar is generated using *CTA macros*, modified by the author of this work to properly estimate the sensi-

tivity for pulsed emission (with a duty cycle of 10%). Total emission (P1 + P2) and both P1 and P2 peaks were simulated using the MAGIC power-law fits given in [27]. CTA candidate layouts “B”, “E” and “C” were tested. Fig. 4.8 shows the estimated Crab pulsar spectrum using Prod-1 configuration “B”, where both total signal and resolved peaks are well characterized.

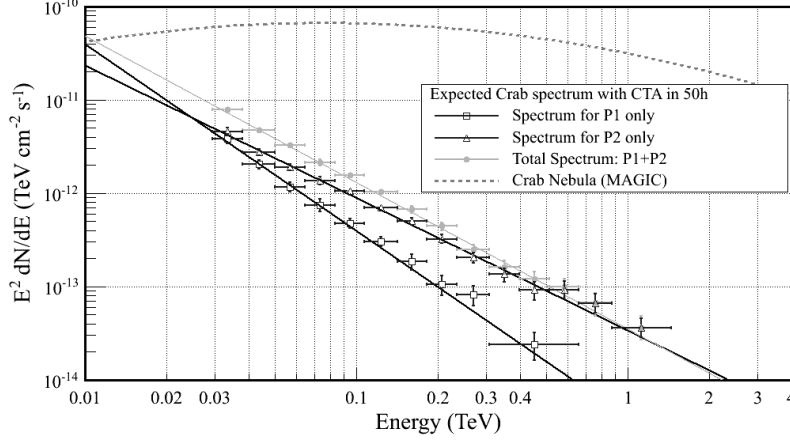


Figure 4.8: Simulated Crab pulsar SED within 50 h of observation with the CTA Prod-1 configuration B. Each of the two Crab phase peaks and the total spectrum are shown, using MAGIC power-law fits given in [27]. Generated using a modified version of *CTAmacros*.

Results show that the CTA potential for pulsar detection seems encouraging. It will be able to reveal the extent of the Crab pulsed emission up to at least 1 TeV. In fact, the bare detection of the pulsations would take less than one hour.

To explore the detectability of *Fermi* pulsars in the power-law scenario, their spectra are extended above the cutoff energy with a power-law tail that assumes the same spectral index (β) as the one found for the Crab, when a broken power-law is applied to fit both *Fermi*-LAT and VERITAS detections, i.e $\beta = 3.52$ [26]. The final extrapolated spectral shapes for 3 out of the 46 pulsars considered are shown in Fig. 4.7.

As described in Sec. 4.1.3, a 90% background reduction is considered assuming a pulsed duty cycle of 10%, systematic errors of 5% and standard detectability conditions ($S > 5\sigma$ in 50 hours of observation time). No gamma-ray emission from a pulsar wind nebula was considered.

4. CTA FORECAST

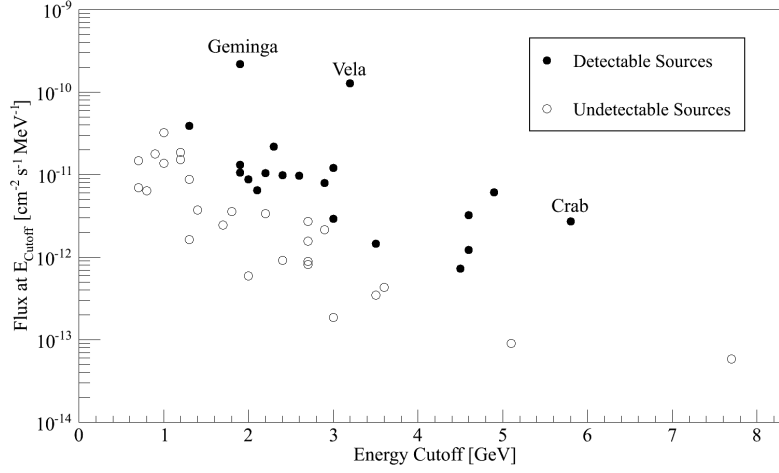


Figure 4.9: Pulsars detectable by the future CTA in 50 hours of observation time assuming their spectra are extrapolated with the Crab pulsar power-law index ($G=3.57$, from [27]). Prod-1 layout “B” was used.

With such hypothetical (except for the case of the Crab pulsar) additional power-law tails all 46 pulsars were then considered as targets for 50 h observations with the CTA configurations: “B”, “C” and “E”. We found that 20 pulsars would be detected with the configurations “B” and “E”. This number reduces to 12 for the configuration “C”. This indicates that configurations “B” and “E” are better suited for pulsar studies than “C” (due to the higher number of LSTs). Fig. 4.9 shows how the detectability with configuration “B” depends on the exponential cutoff energy value (as determined by *Fermi*-LAT) and the photon flux density at this energy. In conclusion, it seems that under the hypothesis of the existence of the VHE Crab-like energy tails, a large fraction (up to $\sim 40\%$ for configuration “B” and “E”) of the presently known brightest *Fermi* pulsars might be detectable with CTA.

On a second step, the possibility of power-law tails with different slopes was investigated. To test such cases, broken power-law spectral shapes were used, in the form proposed by VERITAS [26]. The key parameters in this form are: the break energy E_0 , the slope α of the photon flux spectrum in the *Fermi*-LAT range well below E_0 and the slope β of the photon flux spectrum in the VHE tail, i.e. well above E_0 . For all 46 pulsars the values of E_0 and α were used based on the 1FGL.

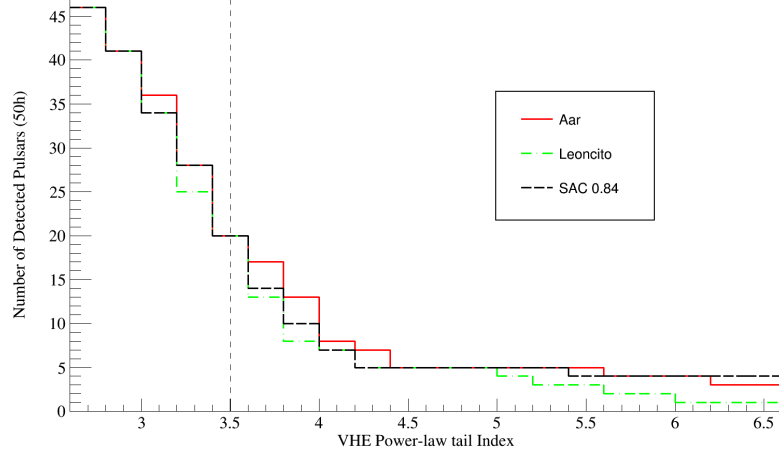


Figure 4.10: Pulsars detectable by the Prod-2 “2A” layout in 50 hours assuming their spectra are extrapolated with a variable power-law tail index. Several site locations of different altitudes were tested: “Aar” (1640 m), “Leoncito” (2662 m) and “SAC” (3600 m).

Using the previous conditions, detectability was tested for Prod-1 configurations. As expected, configuration “B” is the optimal one for all the possible values of β ; the second best is configuration “E” and the worst one is “C”.

To study the impact of the construction altitude, Prod-2 “2A” layout was also used for the 3 sites previously introduced. The studied β values, shown in Fig. 4.10, range between 2 (very hard VHE emission) and 6.5 (very soft VHE emission, almost identical to the exponential cut-off from the 1FGL). High construction altitudes would allow the detection of more pulsars if their emission was consistent with an exponential cut-off, while lower site altitudes would increase the detected VHE γ -ray pulsars population if they had behaviours similar to the one of the Crab pulsar in the VHE range.

Needless to say, there is no assurance that γ -ray pulsars will cooperate in the way described above. However, some theoretical models of young and energetic pulsars as well as old millisecond pulsars speak in favor of pulsed spectral components located in the VHE domain [158]. CTA will be the only facility in the near future capable of solving this problem.

4.2.4 Prospects for Dark Matter searches in dSph galaxies

As introduced in section 1.2.2, WIMP particles with masses between some tens of GeV up to a few TeVs represent one of the most popular scenarios for CDM. Current IACTs have performed several searches for WIMP self-annihilating signals from different sources with expected rich DM content: galaxy clusters [69, 70, 71, 72], the Galactic Center [159], or dSphs galaxies [75, 76, 77, 78, 79, 80]. CTA is planned to perform similar searches, making use of its improved FoV and angular resolution, mainly through deep observations of the Galactic Center and most promising dSph galaxies.

In one hand the Galactic Center seems to be the best option for DM detection, given the close distance and high DM content, although diffuse γ -ray galactic emission is expected, and would hinder an unequivocal DM signature. On the other hand, dSphs may have lower DM content, but they are background free, with approximately known DM density profiles. In this section, prospects of dSph galaxies are explored, assessing the differences between candidate layouts and construction sites in the performance of the CTA for DM detection. For a more general overview of the CTA capabilities on Dark Matter prospects see [23], where part of these results were already published. Note that all dSphs considered in this section (and astrophysical factors) are extracted from reference [23] listed in table 1.1. They have been considered as point-like sources for simplicity (although it may be over optimistic in some cases).

All results that follow assume standard detectability conditions: 5 σ are required for a detection, with the excess above 10 γ -ray events, 3% of systematics (more realistic than the 5% considered in other sections) and $\alpha = 0.2$. The annihilation channels used are the ones present in DAMASCO: $\tau^+\tau^-$, $\mu^+\mu^-$, $b\bar{b}$, and W^+W^- channels from [152], considering 100% branching ratios in each case.

Initially different density profiles are considered, comparing between *cusped* and *cored* ones as in [160], similarly as performed in [23]. The Prod-2 CTA-N layout candidate “2N” simulated at Tenerife and the CTA-S “2A” array at Namibia are compared with the Prod-1 candidate “E”. Fig. 4.11a shows the attained sensitivities on the velocity-averaged annihilation cross-section as a function of the WIMP mass for 100 hours of observation of the best dSph galaxy candidate

in the southern hemisphere, Sculptor, both for NFW and cored isothermal DM halo. CTA-S “2A” candidate array shows a factor ≈ 2 improved upper limits on $\langle\sigma v\rangle$ at a WIMP mass $m_\chi = 1$ TeV with respect to CTA-N “2N” and Prod-1 “E” candidates.

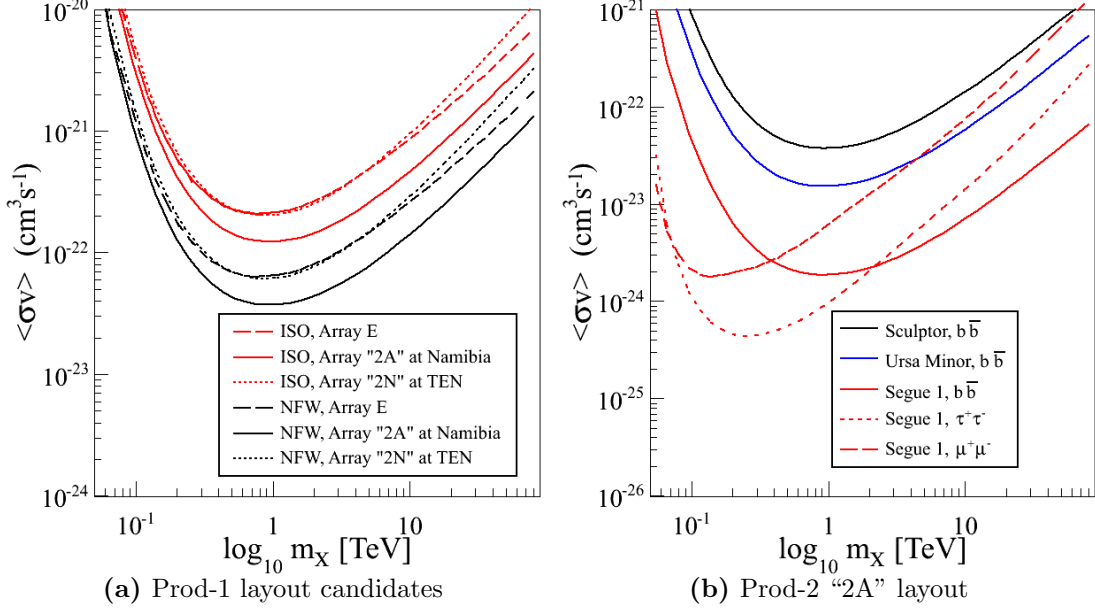


Figure 4.11: CTA sensitivities on the velocity-averaged annihilation cross-section as a function of the WIMP mass for 100 hours of observation for different dSph galaxies and annihilation channels with an integration solid angle (maximizing the astrophysical factor) $\Delta\Omega = 1 \times 10^{-5}$ sr. (*Left*): Sensitivity on Sculptor NFW (black lines) and cored isothermal (ISO, red lines) DM halo profiles by Prod-2 CTA-N layout candidate “2N” simulated at Tenerife, the CTA-S “2A” array at Namibia and the Prod-1 candidate “E”. (*Right*): CTA sensitivity towards Sculptor $b\bar{b}$, Ursa Minor $b\bar{b}$, Segue 1 $b\bar{b}$, Segue 1 $\tau^+\tau^-$ and Segue 1 $\mu^+\mu^-$ channels for CTA-S “2A” layout.

CTA-S “2A” candidate sensitivity in 100 hours is shown in Fig. 4.11b for pointing towards the most promising dSph galaxies both for the Northern and Southern Hemisphere, considering specific annihilation channels with 100% branching ratios: Sculptor $b\bar{b}$, Ursa Minor $b\bar{b}$, Segue 1 $b\bar{b}$, Segue 1 $\tau^+\tau^-$ and Segue 1 $\mu^+\mu^-$. Best constraints correspond to Segue 1, specifically to the $\tau^+\tau^-$ channel for a $m_\chi \approx 300$ GeV and to $b\bar{b}$ for $m_\chi \gtrsim 2$ TeV, more than an order of magnitude away from other dSphs considered.

These channels are selected to calculate the bounds which CTA will be able

4. CTA FORECAST

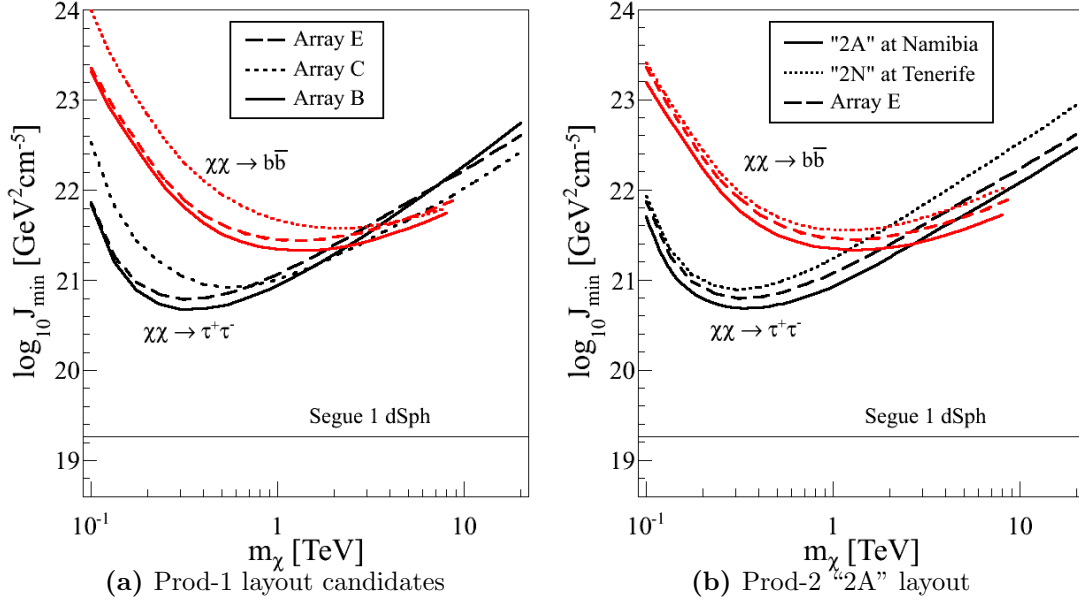


Figure 4.12: Minimum value of the astrophysical factor (J) required for a 5σ detection within 100 hours of observation as a function of the WIMP mass. Two annihilation channels are considered within their valid WIMP mass range: $b\bar{b}$ (red) and $\tau^+\tau^-$ (black). For reference, Segue 1 estimated astrophysical factor is shown with a grey line. (*Left*): Prod-1 candidate layouts “B”, “E” and “C” are considered. (*Right*): Prod-2 candidates CTA-N “2N” simulated at Tenerife and the CTA-S “2A” at Namibia are used along with the Prod-1 “E” layout, for comparison.

to reach on astrophysical factors. The minimum value of the astrophysical factor (J_{min}) required for a 5σ detection within 100 hours of observation as a function of the WIMP mass is shown in figures 4.12a and 4.12b, for several Prod-1 and Prod-2 layouts respectively. Similar limits are attained both for Prod-1 layout “B” (with 5 central LSTs) and Prod-2 layout “2A” (with 4 central LSTs). This effect arises from the better QEs considered in Prod-2, closer to the real efficiencies of current PMTs, which improve low energy performance for Prod-2 layouts. CTA-N prospects on dSphs, reaching roughly 80% worse limits than standard CTA-S “2A”, may be of critical importance for CTA DM prospects as Segue 1, currently the most promising dwarf galaxy (although with larger uncertainties), will be mostly accessible from the northern hemisphere. Nevertheless, the astrophysical factor of the dSph needs to exceed $10^{21} \text{ GeV}^2 \text{ cm}^{-5}$ in order to be detected, which is only 1 or 2 orders of magnitude away than the limits obtained in the Galactic Centre in

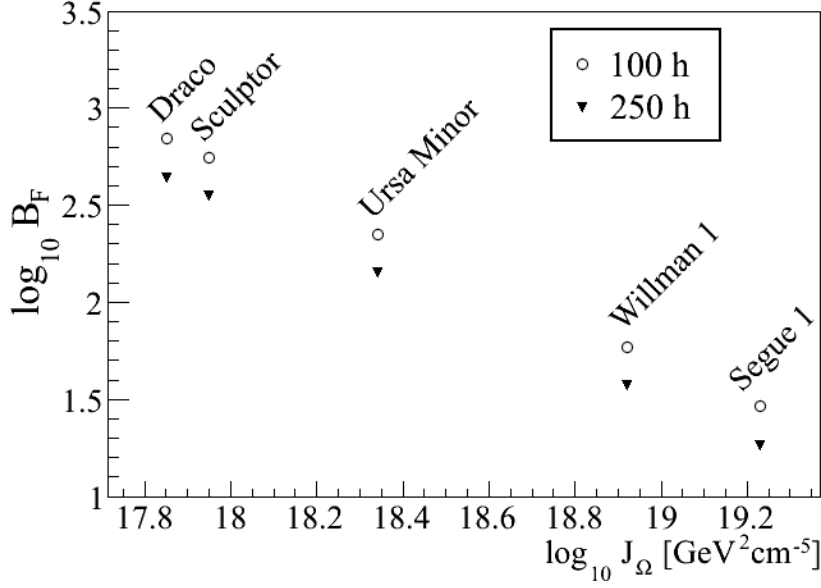


Figure 4.13: Minimum boost factor required for a 5σ detection in 100 h and 250 h by CTA-S “2A” at the Namibian site for the known dSph shown in Table 1.1 from [23] and a 300 GeV WIMP annihilating into $\tau^+\tau^-$.

[23]. Note alternative site altitudes were also tested (“Aar” at 1640 m, “Leoncito” at 2662 m and “SAC” at 3600 m), and differences are less than the 20% in the detectable astrophysical factors, being the Namibian site the one with the best prospects.

In order to illustrate CTA capabilities further, an alternative evaluation of the DM prospects was performed. As briefly introduced in section 1.2.2, the intrinsic flux of dSphs also depends on the *boost factor*, B_F , which stands for inhomogeneities within the DM density profiles, translating into an augmented J factor. The minimum B_F required for a detection of each dSph considered is calculated, taking into account the minimum J factor from Fig. 4.12b for 100 and 250 hours of observation. Again, branching ratios of a 100% are considered for each channel, assuming as the thermal cross-section $3 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$. Fig. 4.13 shows these values for a 300 GeV WIMP particle annihilating into $\tau^+\tau^-$, for 100 and 250 hours of observation considering CTA-S “2A” layout simulated at the Namibian site. Note this values ($\log_{10} B_F > 1$) are not expected from sub-halos in dSph galaxies, which are predicted to range between $0 < \log_{10} B_F < 1$.

4. CTA FORECAST

Chapter 5

Applications of Machine Learning Algorithms to the High Energy Astrophysics

In previous sections, machine learning algorithms have been mentioned in passing. In this chapter the applications of such tools within gamma-ray astronomy is explained, outlining the basic aspects of performance measures, and introduce new implementations for the *Fermi*-LAT and the CTA. In section 5.1 a brief introduction to machine learning algorithms is given, with an special focus on the algorithms applied in the following sections: In Sec. 5.2 MVA methods are applied to CTA MC analysis and compared with alternative methods and in Sec. 5.3 they are applied to determine the source type of 2FGL sources.

5.1 Introduction to Machine Learning Algorithms

Machine learning is a whole scientific discipline consisting on the development of algorithms capable of learning from a data set. These algorithms attempt to find patterns from provided inputs data-driven and subsequently using these patterns to perform predictions. They may be the most extended subfield of artificial intelligence machine learning, and has been applied in every field of science. Some remarkable examples are the classification of DNA sequences [161, 162, 163], the

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

improvement of medical diagnosis [164, 165] and within the astro-particle physics, enhancing the background suppression and event reconstruction in all kinds of experiments [166, 167, 168, 169].

Machine learning algorithms come in two flavours: *Supervised* learning algorithms use labeled datasets with different training parameters or classes. *Unsupervised* algorithms need no pre-defined labels and aim to separate the provided datasets into un-biased subclasses. Given the similar nature of signal and background events within γ -ray astronomy, almost all machine learning methods applied belong to the supervised algorithms class.

Depending on the desired task and output type, supervised algorithms may be applied for *classification*, tagging every prediction to an specific pre-defined category (for instance γ /hadron separation) or for *regression*, providing a prediction of continuous output (such as energy estimation).

Classification is performed by defining boundaries (decision threshold (DT)) of the calculated predictions corresponding to the different classes. These are fixed by searching the values that maximize efficiency on the test samples.

In order to optimize machine learning algorithms and compare different methods, their performance needs to be measured. Performance measurements are typically dependent on the desired task to perform. In the case of binary classifiers the straight forward measurement is the classification accuracy, defined as the percentage of the correct predictions over the total test sample. This measurement may be too simplistic, as information on the amount of false positives/negatives with respect to each class type and specially the distribution of predictions over different DTs may turn out to be useful.

In this chapter two different methods will be applied to measure machine learning algorithms performance:

- Receiver operating characteristic (ROC) curve: Represents the performance of a binary classifier along the different values of the discrimination threshold. It corresponds to the different values of true positive rate and false positive rate obtained for each different DT. As different algorithms may be more efficient at different threshold values, the integral of this curve is usually used to assess machine learning algorithms performance.

-
- Q-value: Is defined as the quotient between the true positives and the square root of the false positives of the test sample. This value is commonly used to measure γ /hadron separation performance, and is defined as $Q = \epsilon_\gamma / \sqrt{\epsilon_h}$, where ϵ_γ and ϵ_h are the resulting efficiencies of surviving γ -rays and hadrons.

Next two machine learning algorithms are briefly introduced: Sec. 5.1.1 describes the RF classifier and Sec. 5.1.2 introduces the Support vector machines (SVM).

5.1.1 Random Forest

Random Forest (RF) is an ensemble classifier that grows a large forest of classification trees that independently make class estimation [170]. Each decision tree selects a number of random input features and creates the best split based on an out of bag (*oob*) random selected set of the whole training data sample. Once the decision forest is built, decision thresholds are computed by counting the votes after running the *oob* datasets through every tree.

A RF classifier is ideal for data mining and variable selection as it incorporates efficient ways of calculating feature importance in the training set. This is achieved by replacing features across classification trees with random values and quantifying the effect of the changes. If the result of the decisions does not change significantly after these changes, the feature has a relatively low importance. On the other hand, if the accuracy rates change dramatically, a particular feature is deemed as important. There is no need for cross-validation with a separate testing set as the process itself computes accuracy rates internally.

In addition to the implementation of the RF present in MARS, in this work the *randomForest* package [171] is used, which adapts the original Random Forests [170] for classification and regression to the R language. The *randomForest* package provides excellent macros for plotting and tuning algorithm parameters. Alternative implementations to this algorithm such as the Boosted Decision Trees (BDT) [172] are used, all within the ROOT library Toolkit for Multivariate Data Analysis (TMVA) [173].

5.1.2 Support Vector Machines

Support vector machines (SVM) have proven to be one of the most effective supervised learning algorithms for pattern recognition [174, 175]. The underlying rationale behind the algorithm seeks to find the optimal margin classifier by constructing a *separating hyperplane* that divides the training set and maximises the separation between different classes, which can then be used either in classification or regression analysis. The points lying closer to the boundaries of a certain hyperplane are called *support vectors*. The latter determine the minimum distances between the hyperplane and their respective classes, the so called *margin*. The maximisation of the optimal margin is computed by taking into account only these vectors, the most representative points to construct the classifier. Complex separating surfaces can be introduced through the use of *kernel* functions, which transform the problem into a linear one in a higher-dimensional space. *Polynomial*, *gaussian* or *radial plane kernel* functions are often used. SVMs excel in performance handling high-dimensional data that can also incorporate the trade off between training errors and overall margins parametrized by a scaling factor Γ and error penalty C .

Several implementations of the SVM algorithm were used in this work. In Sec. 5.2.2 the implementation within TMVA is used while the analysis presented in Sec. 5.3.2 was performed under the R programming language, adopting the *e1071* package as the interface to *libsvm* [176]. Both options offer fast and efficient SVM applications with automatic parameter tuning and allow the use of different pre-defined *kernel* functions.

5.2 MVA in CTA analysis

As introduced in section 3.3.3, there are several applications of multi-variate analysis (MVA) methods within the CTA analysis used in this work. All these methods apply the RF from [170] to improve energy resolution and γ /hadron separation. Improving these methods would allow a better estimation of CTA real capabilities, and provide guidelines to improve its future sensitivity. Also testing alternative algorithms to perform analogous classification may show improvements directly

applicable to the current generation of IACTs. In the following section, two analogous methods implemented for the energy reconstruction: LUTs¹ and RF. We also measure performance and assess the difference between the classical and multivariate approach. In Sec. 5.2.2, the possibility of improving the γ /hadron separation efficiency is tested by implementing new algorithms within the TMVA library in the CTA MC analysis.

5.2.1 Energy reconstruction

As described in Sec. 3.3.3, there are currently two parallel methods to perform the energy estimation in the CTA MC analysis using MARS framework. The first developed approach apply Look-Up Table (LUT) generated for each telescope type, filled with the energy over *size* of MC γ -ray events as a function of the *impact parameter* and the *maximum height*. An alternative method was implemented to improve the energy resolution of simulated array candidates performance, using a RF performing *regression* over the following parameters: The stereo reconstructed *maximum height* and monoscopic *size*, *impact* parameter, *width*, *length*, *concentration ratio*, $\frac{width*length}{size}$, $dist^2$ and the angle between the positive x-axis and the line projected by the source position and the center of gravity of the image.

To compare both methods, equal data samples from CTA-S candidate layout “2Q” simulated at the Namibian site are used for the training of the RF and the LUTs. Subsequent steps in the analysis are performed in the same manner, as a difference in performance should be ascribed to the different method used in the energy reconstruction. An equivalent procedure is performed using MC files from standard CTA-N “2N” layout simulated at the Spanish site location.

As shown in Fig. 5.1, the energy RF method improves resolution in the whole energy range up to a 50%, with no significant effect on sensitivity nor angular resolution. Similar results are observed comparing the resolution of the CTA-N “2N” layout. Note both methods were trained with enough statistics, with identical training samples, so no further improvement is expected using larger training datasets. Note the energy resolution requirements imposed by the CTA consor-

¹Loop-up tables are simple matrixes of values indexing operations to improve computing performance.

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

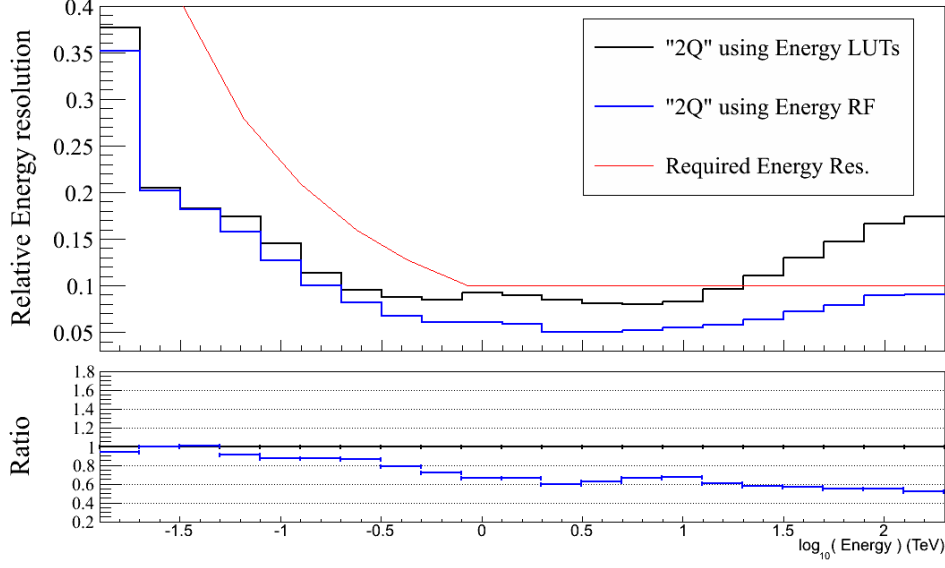


Figure 5.1: Relative energy resolution defined as the 68% containment around $E_{\text{recon}}/E_{\text{true}} = 1$. The *Black* line corresponds to the “2Q” candidate array simulated at the Namibian site, reconstructing the energy using LUTs. The *Blue* line corresponds to the same “2Q” layout simulated at Namibia, using the energy RF. The *red* thin line shows CTA requirements on the energy resolution.

tium are not fulfilled above 1 TeV using the LUT method for the reconstruction, while they are comfortably fulfilled in the whole energy range using the energy RF.

Fig. 5.2 shows the energy reconstructed over the true energy as a function of the estimated energy. A similar averaged bias in the energy is added with both methods, but the energy RF significantly reduces the spread in the distribution returning more accurate estimations.

The *mean gini decrease* is the value calculated by the RF to measure each feature importance in the regression process. Taking a look at the *mean gini decrease* from different telescope types, shown in Fig. 5.3, the reason behind the improvement can be understood. The energy reconstruction through LUTs only used *maximum height* and *impact parameter* to estimate the energy, while RF is able to take advantage of a larger number of variables which greatly affect the performance. *Maximum height* shows low importance, so the LUT is relying on

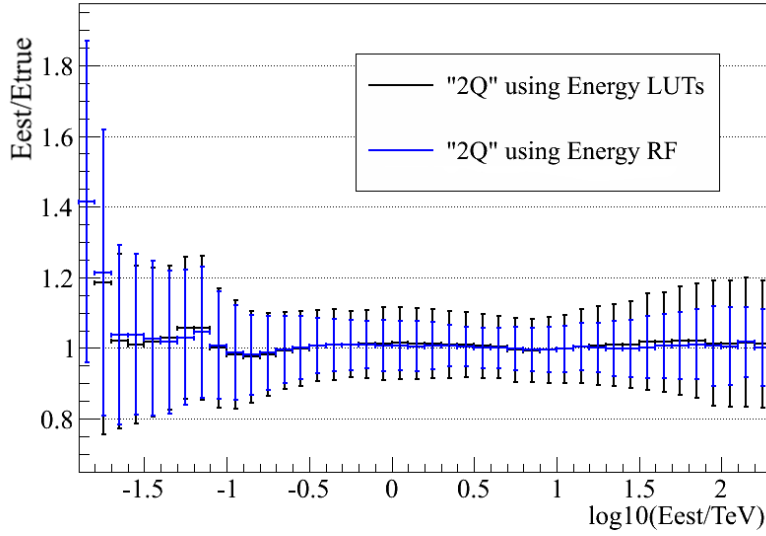


Figure 5.2: Reconstructed energy over estimated energy as a function of the estimated energy. The *Black* line corresponds to the “2Q” candidate array simulated at the Namibian site, reconstructing the energy with the LUT methods. The *Blue* line corresponds to the same “2Q” layout simulated at Namibia, using the energy RF method.

the impact parameter to estimate the energy. Other variables as the $\log_{10}(\text{Size})$, *concentratio ratio* or Dist^2 show as much importance (or event gre ater) than the impact parameter, information exclusively exploited by the RF method.

5.2.2 Gamma-Hadron separation

As RF was created more than a decade ago, it seems plausible that other algorithms may improve the attained efficiencies. TMVA was implemented to work within MARS analysis framework, providing a whole variety of different machine learning algorithms to be applied on CTA MC data. The CTA analysis used in this work was modified to use TMVA libraries instead of those available within the MARS package to perform the gamma-hadron separation, evaluating the *hadroness* parameter. With current modifications, TMVA libraries could be implemented within MARS analysis framework, but that was not within the scope of this work.

Once the alternative method was correctly implemented, different machine learning algorithms were tested:

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

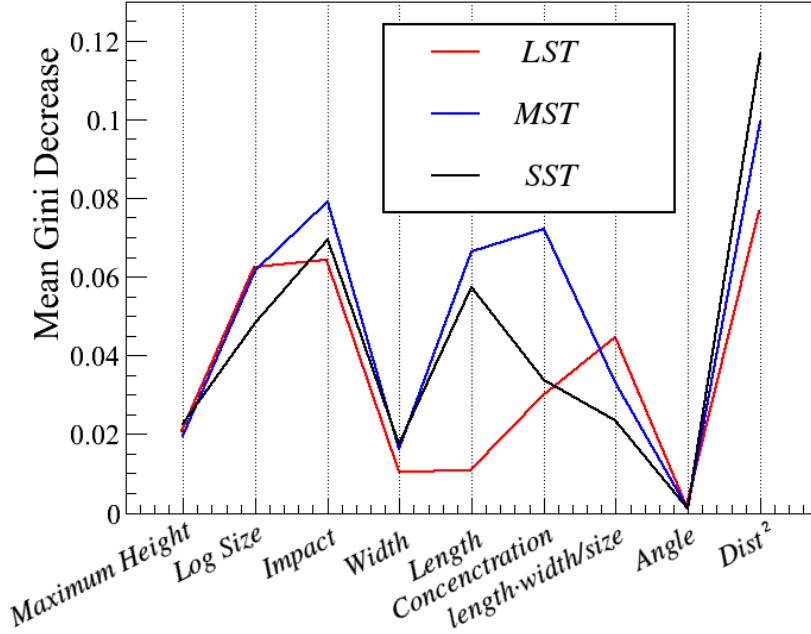


Figure 5.3: Mean Gini decrease calculated by the RF of each different telescope type. The variables are the following: *maximum height*, $\log_{10}(\text{Size})$, *impact* parameter, *width*, *length*, *concentration ratio*, $\text{width} * \text{length} / \text{size}$, the angle between the positive x-axis and the line projected by the source position and the center of gravity of the image and dist^2

- Boosted Decision Trees (BDT): Alternative to RF, also growing a forest of decision trees. But individual trees are not independent, weights are applied consecutively using subsets of trees. Once the forest is created, decision trees are boosted (typically many times), giving larger weights to signal events from the training sample. Different boost methods were tested: *Gradient* (using different bagging fractions), *AdaBoost*, *Bagging* and several combinations from the latter.
- Artificial neural networks (ANN): Similarly as the genetic algorithms, they are inspired by biological processes seen in nature. It makes use of a system of connected “neurons” recursively weighted and transformed by different functions. Two different ANN implementations were tested: The TMLpANN, implementation in ROOT inspired from the MLPfit package by Jerome Schwindling, and the CFMLpANN, by Clermont-Ferrand.

-
- SVM: Described in Sec. 5.1.2. The following kernels were tested: linear, polynomial, Gaussian or sigmoidal.

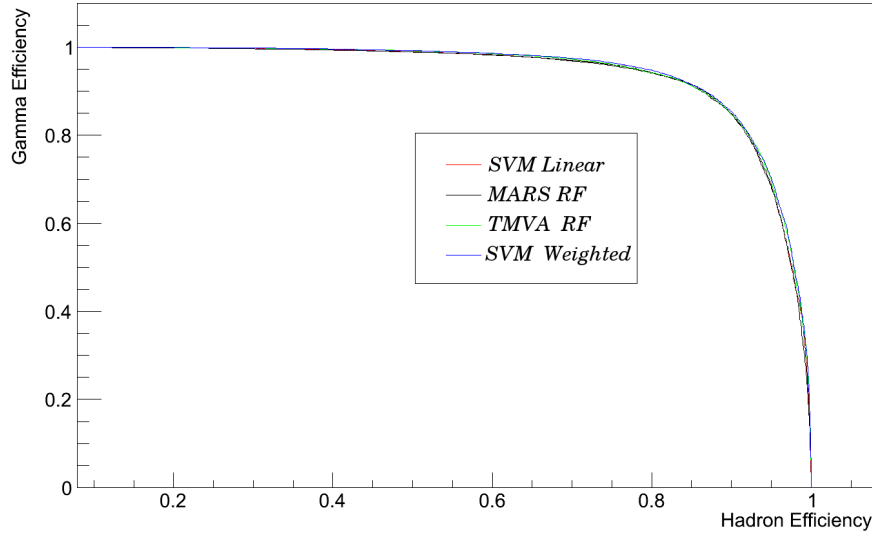
Machine learning algorithms need to be finely tuned in order to significantly improve the classification efficiency. Each tested algorithm was automatically optimized using TMVA methods, maximizing the integral of their ROC curves (described in Sec. 5.1). Initially, a higher number of methods were applied to the training sample. After this unpolished automatic optimization, few algorithms were selected considering their accuracy: TMVA purely randomized BDTs (equivalent to Breiman's RF) and two different configurations of the SVM.

As described in the introduction, there are several accepted practices to compare the performance of binary classifiers. Here, we used ROC curves generated from the selected MVA methods, trained with equal samples and then applied to identical test samples. The study was performed using CTA MC data of a reduced layout, composed by the 4 central LSTs present in most candidate arrays. Note the reduced amount of proton statistics make classification results unreliable at high energies, consequently just energies below 1 TeV were considered. As pictured in Fig. 5.4 all methods show good classification, although some differences appear with a closer look 5.4b.

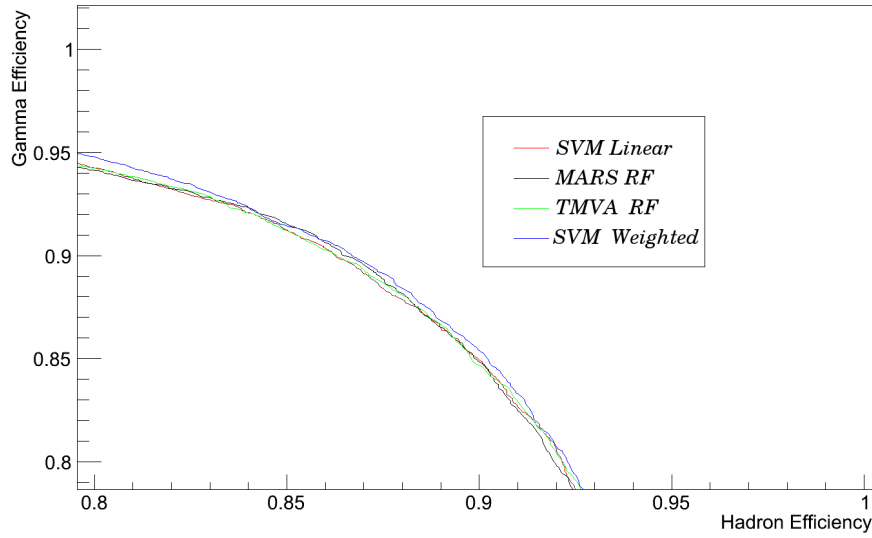
Note these curves were created from the whole energy range, tuning the MVA methods with no further selection. As the statistics are dominated by low energy events, parameter tuning may turn out to be too focused on improving low energy events classification, as these curves make no distinction between different energy ranges. These classification algorithms could be improved by tuning different methods (with different parameters) as a function of the energy, or weight the events assigning more importance to the higher energy events, compensating their low frequencies.

To test the separation power as a function of the energy, the decision threshold maximizing the *quality* factor was selected for each energy bin considered, defined as $Q = \epsilon_\gamma / \sqrt{\epsilon_h}$. Fig. 5.5 shows the obtained Q values as a function of the energy, showing all methods follow a similar trend, and no significant differences are observed. This result seems to show that upgrading the algorithm may bring slight improvement to overall sensitivity. Instead of changing the applied algorithm,

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS



(a) ROC: full efficiencies range



(b) ROC: region of interest

Figure 5.4: ROC curves of the binary classifiers used for γ -hadron separation: default RF implementation within the MARS package, SVM with linear kernel and default tuning, SVM with manually weighted parameters and TMVA BDTs equivalent to the RF. *Up:* Whole range of efficiencies, showing a very similar performance of all tested methods. *Down:* Same curves over the range of efficiencies between 0.8 and 1, to show slight differences between tested models.

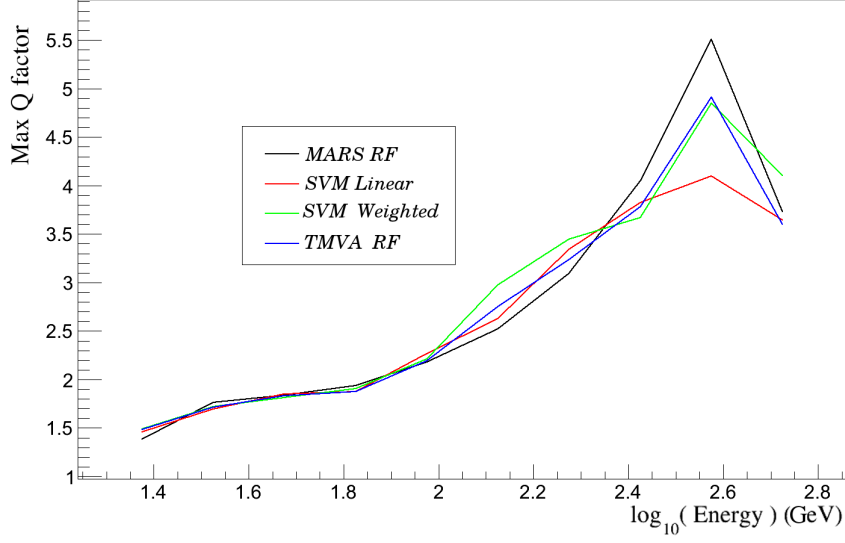


Figure 5.5: Q factors as a function of the energy for the different algorithms tested: default RF implementation within the MARS package, SVM with linear kernel and default tuning, SVM with manually weighted parameters and TMVA BDTs equivalent to the RF. Datasets correspond to Prod-1 data corresponding to a reduced layout of the 4 central LSTs. Note the reduced amount of proton statistics make classification results unreliable at higher energies, so just energies below 1 TeV were considered.

adding new variables carrying information of the shower geometry could lead to improvements in this area.

5.3 MVA for source type determination in 2FGL

As introduced in Sec. 2.3.1, the extraordinary success of the *Fermi* mission marks the beginning of the golden age for γ -ray astrophysics. With 24 months of data, the Second *Fermi* LAT Catalogue (2FGL) listed 1873 sources in the 100 MeV to 100 GeV energy range, of which 1092 objects were connected with known AGN at other wavelengths and 108 were pulsars. While *Fermi* has greatly mitigated issues inherent to source localisation in the γ -ray regime, 269 sources in the 2FGL (15% of the total) remain without obvious counterparts at Galactic latitude $|b| \geq 10^\circ$. In addition, source determination lacks multi-wavelength counterparts to characterize AGN subclasses, and 257 extragalactic sources are of uncertain type.

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

Such failure to associate the entire *Fermi* catalogue continues to fuel speculation about the existence of new types of gamma-ray source classes. In Sec. 5.3.1 MVA methods are applied as an attempt to unveil the nature of unassociated *Fermi* sources investigating the possibility of identifying dark matter subhalo candidates. In Sec. 5.3.2, in order to understand all the intricacies of the AGN population and guide multi-wavelength observations, the possibility to predict specific AGN subclasses based on the observed γ -ray spectral properties is explored.

5.3.1 Sybil

Probably the most intriguing potential sources of γ -ray emission are the dark matter subhaloes [177, 178]. Numerical CDM simulations suggest that galaxies like our own are surrounded by a wealth of small dark matter subhaloes that survived structure formation [179, 180]. Massive subhaloes ($M \geq 10^7 M_\odot$) would correspond to “classical” dwarf galaxies. Less massive ones would be optically elusive and might only be revealed as γ -ray point sources when WIMP annihilate to gamma rays [181]. As a result, nearby dark matter subhaloes might be lurking among the unassociated *Fermi* sources at high Galactic latitudes. If found, an annihilation signal from Galactic subhaloes would clinch the first non-gravitational signature of dark matter.

The hunt for dark matter subhaloes in the *Fermi* catalogue is currently underway [182, 183, 184, 185]. Most approaches involve the hypothesised sharp spectral cut-off or step expected at the WIMP mass [186]. Assuming that the WIMP mass falls between 100 MeV and 50 GeV, a dark matter subhalo could be detectable in the *Fermi* MeV-GeV band, but would disappear in the GeV-TeV band, effectively creating a TeV dropout.

Here the possibility of identifying dark matter subhalo candidates is investigated using supervised machine learning algorithms. Rather than starting with an *ad hoc* theoretical dark matter spectrum it would be desirable to exploit pattern recognition of known γ -ray features in associated sources and use this information to locate outliers that might constitute novel emitters. Machine learning algorithms have already been used to study the 1FGL. For example, [187] investigated classification trees and logistic regression to predict classes of unassociated sources

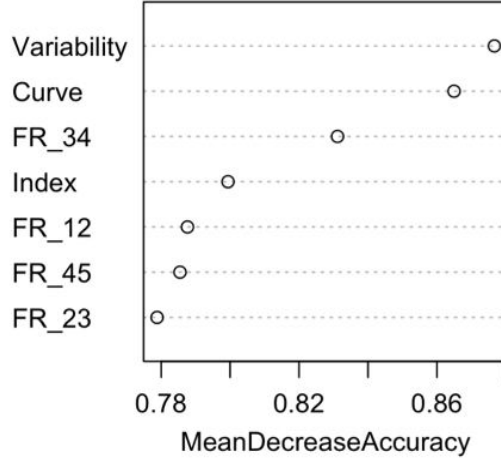


Figure 5.6: γ -ray features ranked in order of importance. **MeanDecreaseAccuracy** measures the difference between accuracy rates before and after permutation of individual features averaged over all trees. Higher percentages indicate more importance.

in the 1FGL based on a set of γ -ray features. K-means clustering was also applied to help distinguish individual counterparts within *Fermi* error contours [188].

5.3.1.1 Datasets

To prepare the required dataset, the complete *Fermi*-LAT 2FGL catalogue was collected, consisting of 1873 sources (100 MeV-100 GeV) of which 1300 are firmly identified/associated and 573 are unassociated sources [109, 189]. In total, a list that includes 800 labelled AGNs (BL Lacs and FSRQ only) and 108 pulsars was considered. There are additional γ -ray classes in the 2FGL, but AGNs and pulsars are the largest and most common at $|b| \geq 10^\circ$. Thereby a simple bimodality of classes was considered. For each of the 908 sources a total of 68 features are reported in the 2FGL. Features include Galactic latitude, Galactic longitude, spectral index (Index), variability, curvature index (Curve), and fluxes in five bands. In addition, four derived features were generated defined by flux ratios $FR_{ij} = Flux_i / Flux_j$ between consecutive bands for 0.1–0.3 GeV (Band 1), 0.3–1 GeV (Band 2), 1–3 GeV (Band 3), 3–10 GeV (Band 4), and 10–100 GeV (Band 5) comparable to the features first introduced by [187].

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

To avoid working with too many features that could generate noise in the classifier, it was first performed an identification of the subset of features that best discriminates what constitutes an AGN or a pulsar. For that purpose, the relevance of each feature was computed towards the target class, rank them by importance, and apply the classifier to a subset of the most relevant ones. Specifically, the measure of importance **MeanDecreaseAccuracy** was used, implemented within `randomForest` [170, 190]. Initially, the accuracy rate is computed for each tree as the Random Forest is constructed. The value of a particular feature is then permuted across all the objects while other features are left unchanged and the accuracy rate is recorded again. The **MeanDecreaseAccuracy** is the overall percentage decrease in accuracy rate averaged over all trees. If the feature is important, there should be a greater decrease in the accuracy rate compared to the initial one. Figure 5.6 shows the top most important features ranked by importance. The features that most clearly differentiate AGN and pulsar classes include: Index, Curve, Variability, and Flux Ratios (FR_{12} , FR_{23} , FR_{34} , and FR_{45}). This selection is in general agreement with [187] who chose similar features for supervised classification of the 1FGL. Additional features showed considerably smaller values in their importance (**MeanDecreaseAccuracy**) and are thus not considered in the analysis.

In order to construct and train *Sibyl*¹, the available dataset consists of 800 AGNs and 108 pulsars. However, given the highly imbalanced nature of the sets, the pulsar sample is replicated to attain a closer size as the AGN class [191, 192]. Practically, the content of the datasets have not changed but the replication mechanism adds weight to the minority sample and achieves improved performance in the classifier.

After matching the AGNs and pulsar datasets, 100 alternate training and testing sets are generated built from randomly selected objects (2/3 and 1/3 of the sample respectively). They are then used to produce Random Forest models with 500 trees for each training set. For validation, individual performance is evaluated at each of the 100 testing sets. Accuracy rates are computed directly by comparing the class predicted by *Sibyl* with the true class for each object in the testing

¹In ancient Greece, a sibyl was a person or agency considered to be a source of predictions or oracles.

sets. On average, an accuracy rate of 97.1% based is achieved on majority voting (97.7% for AGNs and 96.5% for pulsars). Inclusion of absolute Galactic latitude $|b|$ in the classifier lowered AGN and pulsar accuracy rates slightly to 97.4% to 95.5% respectively. Since pulsars tend to be situated along the Galactic plane and AGN are more numerous at high Galactic latitude, it is possible that using Galactic latitude as a feature could introduce a tiny bias against AGN near the Galactic plane and pulsars away from it [187]. Generally, most of the misclassifications occur when less than 70% of the individual trees ($P < 0.7$) agree on a classification. Figure 5.7 displays the outstanding separation between AGNs and pulsars, which explains the high accuracy rates obtained by *Sibyl*.

5.3.1.2 Application to unassociated sources

The designation of 2FGL sources usually falls into three categories: identified, associated, and unassociated. A firm identification of a γ -ray source can only be established through contemporaneous temporal variability, similar spatial morphology, or equivalent pulsation at other wavelengths. An association only requires positional correspondence of a possible counterpart with a 2FGL source. Unassociated sources lack a formal counterpart at other wavelengths.

Here, a fourth category is considered to designate 2FGL sources: “prediction”. The objective is twofold: to predict the class of high-latitude unassociated *Fermi* objects in the 2FGL; and to produce a list of outliers that could be explored as potential dark matter subhaloes. For each of the 269 unassociated *Fermi* sources at $|b| \geq 10^\circ$, *Sibyl* provides a prediction that the object is an AGN (P_{AGN}) or a pulsar (P_{Pulsar}) based on individual votes polled from all trees in the decision forest.

Since the objective is to isolate outliers that might constitute dark matter subhalo candidates, only *Sibyl* predictions with $P > 0.7$ are accepted *i.e.*, at least 70% of the trees agree on the final decision. Otherwise, the object remains without a prediction. Such threshold value is set based on the results explained in Section 5.3.1.1. In total, *Sibyl* predicts 216 objects to be AGN and 16 to be pulsars. The resulting predictions and percentages of voting agreements are listed in Table 5.1. Finally, the remaining 37 objects left without a firm prediction are the focus of

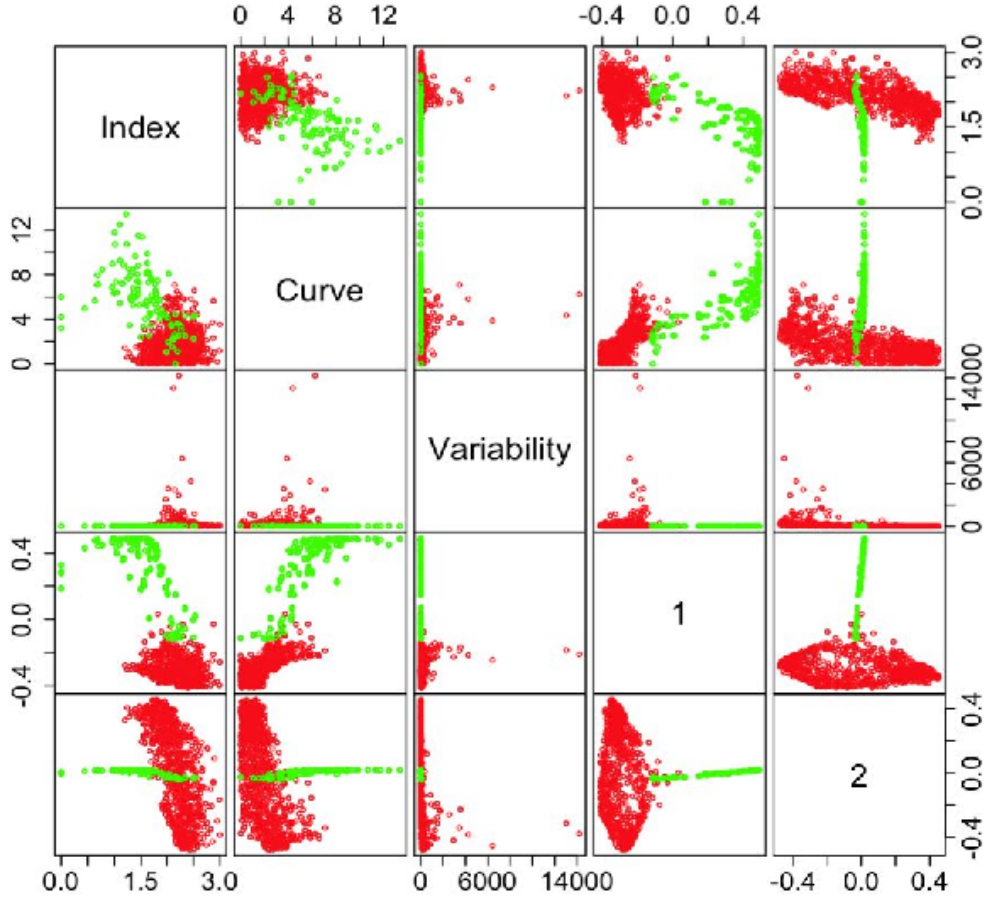


Figure 5.7: Properties of *Fermi* features plotted against each other. Top features include index, curve, variability, and 1st and 2nd scaling coordinates (1 and 2 respectively) generated by *Sibyl*. Two distinct classes are clear: AGNs (red) and pulsars (green).

our outlier study in the next section. It is important to note that under some specific circumstances, dark matter subhaloes could mimic the spectral properties of certain pulsars [184, 193]. This possibility is discussed further in 5.3.1.4.

5.3.1.3 Search for dark matter subhaloes in the 2FGL

In order to better understand the nature of the remaining 37 objects their outlyingness is computed, which is a measure of how far away an object is from its closest class. Apart for predicting an object's class, Random Forest computes the

Table 5.1: Predictions and voting percentages for unassociated *Fermi* sources in the 2FGL, ordered by RA

Source	P _{AGN}	P _{Pulsar}	Prediction
2FGL J0004.2+2208	0.974	0.026	AGN
2FGL J0014.3-0509	0.992	0.008	AGN
2FGL J0031.0+0724	0.946	0.054	AGN
2FGL J0032.7-5521	0.998	0.002	AGN
2FGL J0039.1+4331	0.776	0.224	AGN
2FGL J0048.8-6347	0.922	0.078	AGN
2FGL J0102.2+0943	0.998	0.002	AGN
2FGL J0103.8+1324	0.998	0.002	AGN
2FGL J0106.5+4854	0.406	0.594	—
2FGL J0116.6-6153	0.992	0.008	AGN
2FGL J0124.6-2322	0.998	0.002	AGN
2FGL J0129.4+2618	0.820	0.180	AGN
2FGL J0133.4-4408	0.968	0.032	AGN
2FGL J0143.6-5844	1.000	0.000	AGN
2FGL J0158.4+0107	0.990	0.010	AGN

Note: The complete list of predictions is available at <http://www.gae.ucm.es/~mirabal/sibyl.html>

proximity of each predicted *Fermi* object n to every element k within each class $\sum_{\epsilon_{class}} prox(n, k)$. Formally, each individual proximity $prox(n, k)$ is computed as the fraction of trees in which elements n and k fall in the same terminal node [170, 171]. The outlyingness of an element n is calculated as the reciprocal sum of the squared proximities to all objects within its class. This outlying measure is normalised by subtracting the median and dividing by the absolute deviation from the median [171]. Larger outlyingness values are common in objects that are extremely different from the average, which could correspond to dark matter subhaloes. Figure 5.8 shows the distribution of outlyingness for the 37 objects without firm predictions. For comparison, the outlyingness for the remaining 232 objects that were predicted by *Sibyl* in the previous section is also plotted. Additionally, Table 5.2 lists the five objects with the largest outlyingness.

Given that outlyingness values much greater than 10 usually indicate novel cases [170], there is no strong indication of novelties (significant outliers) among the 37 objects without firm predictions. The top five outliers have an average

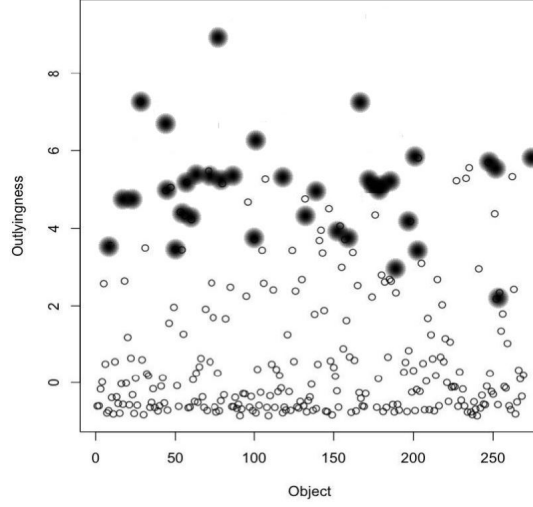


Figure 5.8: Distribution of outlyingness for the 37 objects without firm predictions (shaded circles) and the 232 predicted by *Sibyl* (open circles). The top outliers are summarised in Table 5.2.

Table 5.2: Top outliers among high-latitude unassociated sources in the 2FGL

Source	P_{AGN}	Outlyingness
2FGL J0953.6–1504	0.658	9.0
2FGL J0418.9+6636	0.574	7.2
2FGL J1710.0–0323	0.500	7.1
2FGL J0533.9+6759	0.336	6.6
2FGL J0336.0+7504	0.476	6.2

flux of 1.1×10^{-9} ph cm $^{-2}$ s $^{-1}$ (1–100 GeV) while unassociated source fluxes at high latitudes range from 7.7×10^{-9} to 1.1×10^{-10} ph cm $^{-2}$ s $^{-1}$. Thus, they are not necessarily the faintest sources in the dataset. On the other hand, the mean photon index of sources in Table 5.2 is 2.2 ± 0.3 , while photon indices in the unassociated sample range from 1.1 to about 3.0. Inspection of individual features in this manner yields limited insight into what makes these outliers stand out from the rest of the sample. As mentioned before, the exploration of the entire feature space is precisely where the supervised learning algorithm excels. Unfortunately, *Sibyl* cannot assess by itself whether the outlyingness is due to an anomaly in the data taking process, a simple variation within known *Fermi* classes, or a true novel

source class such as dark matter annihilation in Galactic subhaloes.

5.3.1.4 Discussion

The results show that machine learning algorithms provide a reasonable route not only to predict unassociated AGN/pulsars in the 2FGL, but also to produce a list of sources with unusual features that could be explored as potential dark matter subhalo candidates. After training on 908 identified/associated *Fermi* objects, *Sibyl* has been applied to predict the class of unassociated *Fermi* sources in the 2FGL. Out of 269 unassociated sources at high latitudes, 216 are found to be AGN candidates and 16 are considered potential pulsars with prediction accuracy rates greater than 96.5%. *Sibyl* has also produced a list of 37 outlier objects; however, none of these exhibits significant outlyingness that can be directly connected to new γ -ray classes (including dark matter subhaloes) at this point. Note these results are strict predictions based on pattern recognition and thus a rigorous source identification process will have to localise actual counterparts at other wavelengths.

The results leave some room, albeit very small, to accommodate dark matter subhaloes or alternative source classes in the 2FGL. These pockets could be targeted to exhaust all possibilities. Looking forward, zooming in on a reduced group of sources might be a wise observational strategy. For obvious reasons, the set of objects with the largest outlyingness could be a reasonable place to conduct a dedicated survey. If dark matter consists of particles with a mass below 60 GeV [194], dark matter subhaloes might also be camouflaging among the ranks of predicted pulsars as their spectral signature could be similar to the pronounced spectral cut-off predicted by certain dark matter models. However, a number of these sources could be old radio-quiet pulsars which will complicate the search for a counterpart [195].

There are a number of issues that need further exploration. For instance, the predictions are heavily dependent on the robustness of the spectral parameters listed in the 2FGL. Most machine learning algorithms lack a proper treatment of uncertainties in each of the features considered [196, 197]. Inclusion of uncertainties as individual features in *Sibyl* did not yield improved performances in our predictions. With additional years of flight, *Fermi* will likely keep improving the

accuracy of the γ -ray features. However, attempts should be made to account for feature errors properly.

Ultimately, the main reason that a large *Fermi* fraction remains unassociated to begin with has to do with the quality of localisations in the γ -ray band. At faint flux levels, it becomes ever more difficult to associate a *Fermi* source with a particular counterpart. The best association procedures rely on positional coincidences and correlations with flat-spectrum radio sources [189]. None the less, considering the results presented here and the scatter in γ -ray flux it seems likely that many of the unassociated sources at high latitude are AGNs or mid-latitude pulsars with somewhat fainter radio fluxes than their brighter cousins.

Without a major breakthrough in localisations, the actual counterparts of most unassociated *Fermi* objects will be difficult to pinpoint in the short term. Eventually, a significant improvement in localisations will come, particularly for Galactic sources, courtesy of the future CTA that will achieve enhanced angular resolution within the TeV range, as shown in Fig. 2.8.

5.3.2 AGN type determination

Out of the 1092 sources designated as AGN in the 2FGL, 436 are BL Lac, 370 are FSRQ, 12 are radio galaxies, 6 are Seyferts and 11 are other AGN. Despite this important level of achieved sophistication, the remaining 257 sources are designated as active galaxies of uncertain type (AGU) that total 25% of all AGN. Generally, AGU are positionally coincident with flat-spectrum radio sources showing distinctive broad-band blazar characteristics, but lacking reliable optical measurements [189].

In order to understand all the intricacies of the AGN population, it is important to take further steps to assess the nature and redshift of the sources classified as AGU. In the past, this has been accomplished via a two-step approach. The initial classification of an AGN relies on painstakingly dedicated optical spectroscopy to help identify unique emission or absorption features [198]. If no significant features are found, the second step consists of multi-band photometry to help estimate the redshift of suspected BL Lacs [199, 200, 201].

Without optical spectroscopy, generally there is no sufficient information to

directly determine whether an individual source is a BL Lac or a FSRQ. Unfortunately, optical spectral observations are taxing and can take years to complete. Ideally, one would like to find a discriminator for distinct source subclasses that relies solely on readily available observational characteristics. Recently, [202] introduced a method that helps recognise γ -ray blazar subclass based on infrared colours from the Wide-Field Infrared Survey Explorer (WISE). In this section, the possibility of determining AGN subclass for *Fermi* sources directly from γ -ray spectral features extracted from the 2FGL is explored.

In particular, two supervised machine learning algorithms are employed, Random Forests and Support Vector Machines, initially trained on identified/associated AGN and subsequently used to infer specific blazar subclass of AGN of uncertain type. This section may be considered a natural extension of previous machine-learning strategies introduced to predict source class in unassociated Fermi point sources [187, 203, 204].

5.3.2.1 Datasets

In this case, the trained classifiers need to distinguish between two AGN classes: BL Lacertae (BL Lac) and flat-spectrum radio quasars (FSRQ). In the 2FGL, there are a total set of 1074 identified/associated AGN objects with the following labels: “bzb” (BL Lacs), “bzq” (FSRQs), “agn” (other non-blazar AGN) and “agu” (active galaxies of uncertain type). From this global set, identified/associated blazars (“bzb” and “bzq” labels) are grouped as the training/testing set of the used algorithms. Training dataset is composed of 806 sources, divided in a fairly balanced manner that includes 370 FSRQs and 436 BL Lacs. In addition, undetermined sources (“agu” labels) are placed in a separate dataset consisting of the 257 objects. Once the algorithms are trained and tested, the classifiers are applied to the latter. Note that the initial approach is a simple binary classification problem that attempts to distinguish whether an individual AGU is a BL Lac or a FSRQ. It is possible that other subclasses are represented within the AGU dataset. However, additional AGN subclasses only account for 3% of the whole sample.

The next step involves choosing from the different γ -ray spectral features available for each source. Although the algorithms are not strongly affected by noise,

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

it is relevant to limit misleading features that might affect the characterisation. Initially, all basic features reported in the 2FGL [109] are selected. As in Sec. 5.3.1, Hardness Ratios ($HR_i = Flux_i - Flux_j / Flux_i + Flux_j$) and Flux Ratios ($FR_{ij} = Flux_i / Flux_j$) are supplemented, ending up with a set of 20 distinct features. Armed with this set of variables, feature importance is computed to find those most representative with a robust method already implemented in the *randomForest* package [171, 203]. This process outputs two measures of importance: *MeanDecreaseAccuracy* and *MeanDecreaseGini*. Both are excellent indicators of feature relevance [170].

Once feature importance measures are computed, new sets of data are created with different number of features by iteratively removing the variables with lower *MeanDecreaseAccuracy*, and comparing accuracy rates attained by RF and SVM algorithms on these sets. Although RF does not require a tailored training/testing analysis to estimate accuracy rates, it is useful to compare both algorithms directly with identical training/testing sets. Through feature selection, the initial 20 features is downsized to a final set of 9. The final set of variables includes (ordered by decreasing *MeanDecreaseGini*) **Powerlaw Index** (76.6), **Pivot Energy** (59.2), **Flux Density** (27.1), **Variability Index** (20.1), **Flux1000** (12.6), and four **Hardness Ratios**: HR_2 (19.4), HR_1 (17.5), HR_3 (14.4) & HR_4 (10.6). Features considered but later discarded include **Spectral Index**, **Energy Flux**, **Curvature Index**, **Flux** in five different energy ranges, and **Flux Ratios**.

The top two most representative features for AGN subclass determination are **Powerlaw Index** and **Pivot Energy**. The clean separation between blazars is obvious in Fig. 5.9 and it intuitively stands on observational arguments. As explained in [189], there is a well established spectral difference in the LAT energy range between FSRQs and BL Lacs. In general, the AGN IC peak is located at lower energies for FSRQs and at higher energies for BL Lac objects. Typical values are 1 MeV – 1 GeV for FSRQ and 100 MeV – 100 GeV for BL Lacs respectively [205].

The overall effect is that FSRQs show softer spectra than BL Lacs, and therefore, higher values of **Powerlaw Index**. **Pivot Energy** is defined as the energy at which the relative uncertainty on the differential flux is minimal. It is also an estimate of the point where the covariance of **Powerlaw Index** and **Flux Density** is

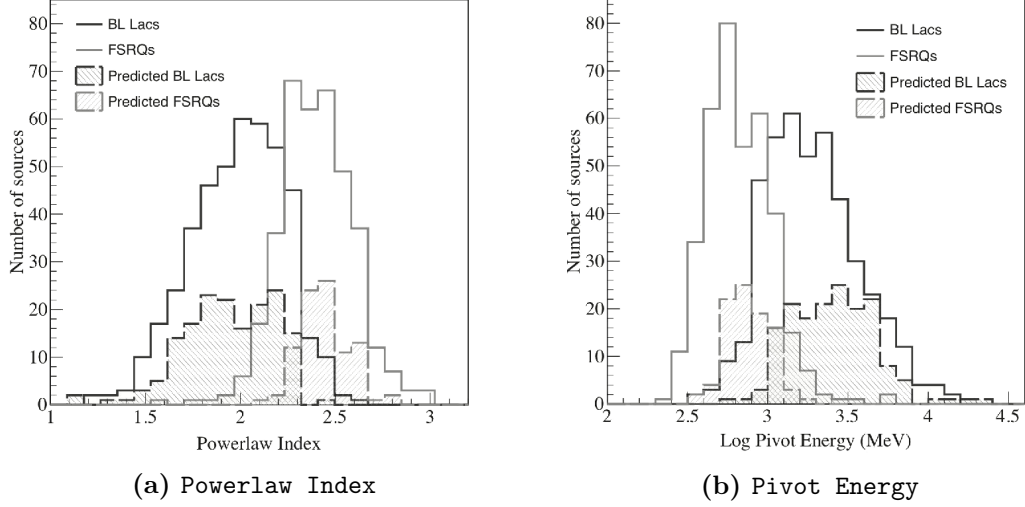


Figure 5.9: Distributions of **Powerlaw Index** (left) and **Pivot Energy** (right) for identified/associated BL Lacs (black) and FSRQs (grey). The filled areas show results for AGU: Predicted BL Lacs (filled dark area) and predicted FSRQs (filled light area).

minimised [109]. The relative dominance of lower energy events for FSRQs places the general location of the **Pivot energy** at lower energies compared to BL Lac spectra. As a result, the difference found in **Pivot Energy** between both populations can be understood as the overall effect of the spectral characteristics of FSRQs and BL Lacs produced by the difference on the position of IC peak in the spectral energy distribution for both populations.

Both SVM and RF algorithms require parameter tuning to achieve their best performance. In the case of SVMs, there is an automatic tuning process *best.tune* that returns the appropriate values of C and γ for a particular *kernel* function and training set. In order to make a selection, the classification accuracies were scanned for different *kernel* functions and used the tuned parameters to discriminate amongst them. *Linear*, *polynomial*, *sigmoid*, and *radial* kernels were tested. For the final training set, a *C-classification linear kernel* with $C = 1$ and $\gamma = 0.11$ was used. For RF, *tuneRF()* performs an automatic search for the most efficient number of features used per classification tree for a chosen training set [171]. Ultimately, 9 spectral features were employed, four variables randomly sampled at each split (node size), and a total of 5000 trees.

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

After culling the datasets with the chosen features and tuning the algorithms for best performance, testing is performed to estimate the error of the resulting classification. As the training set, a random selection of 2/3 of all identified/associated AGN was used and the remaining 1/3 as a the testing set. To estimate the accuracy rates, the actual source class was compared with the class predicted by each classifier. For 500 of these training and testing sets, average accuracy rates of 85% were obtained, adopting a decision threshold of $P > 0.5$ for both RF and SVM. Note that with such threshold there are few ambiguous events since both P_{SVM} and P_{RF} are required to be greater than 0.5. Is a more conservative condition is considered, for instance $P > 0.8$, the accuracy rates improve up to 94%. In this case, there is a bigger fraction of the sample that remains untagged.

For further verification, accuracy rates were calculated by leaving one object out from the training set and using that single object as the testing set. The leave-one-out cross validation rate is 85% for common decision threshold of $P > 0.5$ and 95% with $P > 0.8$ showing that larger training sets do not produce significant increases in accuracy rates.

5.3.2.2 Results

Once the classifiers have been trained and tested, both algorithms are applied to the set of AGN of uncertain type. For each of the 257 AGU, the classifiers return a decision threshold that an individual object is a BL Lac (P_{bzb}) or a FSRQ (P_{bzq}), where $P_{bzq} = 1 - P_{bzb}$. A fraction of the resulting predictions is listed in Table 5.3. Decision thresholds P_{bzb} calculated with both RF and SVMs are shown, as well as a class prediction satisfying the condition $P(RF)$ and $P(SVM) > 0.8$. Fig. 5.10 shows $P(RF)$ and $P(SVM)$ values obtained with each classifier for the 257 sources. Overall, there is an agreement rate of 91% between the algorithms. Though there are some discrepancies (for instance RF show higher BL Lac classification rates than SVMs), the results are outstanding considering the distinct underlying assumptions of the algorithms.

Table 5.4 shows overall numbers sorted according to different criteria imposed for both RF and SVM classification. In particular the predicted number of occurrences is listed in terms of different decision thresholds ($P > 0.5$, 0.8, and

Table 5.3: Predictions for *Fermi* AGN of uncertain type in the 2FGL, ordered by RA. Threshold values $P_{bzb} < 0.2$ (in the case of FSRQs) and $P_{bzb} > 0.8$ (in the case of BL Lacs) must be met in both methods.

Source	P_{bzb} (RF)	P_{bzb} (SVM)	Prediction
2FGL J0001.7-4159	0.84	0.80	bzb
2FGL J0009.1+5030	0.97	0.95	bzb
2FGL J0009.9-3206	0.53	0.57	-
2FGL J0010.5+6556c	0.14	0.07	bzq
2FGL J0018.8-8154	0.69	0.80	-
2FGL J0019.4-5645	0.16	0.04	bzq
2FGL J0022.2-1853	0.99	1.00	bzb
2FGL J0022.3-5141	0.46	0.50	-
2FGL J0038.7-2215	0.99	1.00	bzb
2FGL J0044.7-3702	0.06	0.04	bzq
2FGL J0045.5+1218	0.91	0.85	bzb
2FGL J0051.4-6241	1.00	1.00	bzb
2FGL J0055.0-2454	1.00	0.99	bzb
2FGL J0056.8-2111	0.97	0.99	bzb
2FGL J0059.2-0151	0.95	0.99	bzb
2FGL J0059.7-5700	0.03	0.02	bzq
2FGL J0103.5+5336	0.93	0.94	bzb
2FGL J0110.3+6805	0.86	0.68	-
2FGL J0118.6-4631	0.96	0.98	bzb
2FGL J0127.2+0324	0.98	0.99	bzb
2FGL J0131.1+6121	0.93	0.97	bzb
2FGL J0134.4+2636	0.99	0.97	bzb
2FGL J0137.7+5811	0.44	0.38	-
2FGL J0146.6-5206	0.95	0.92	bzb

Note: The complete list of predictions is available at <http://www.gae.ucm.es/~thassan/agus.html>.

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

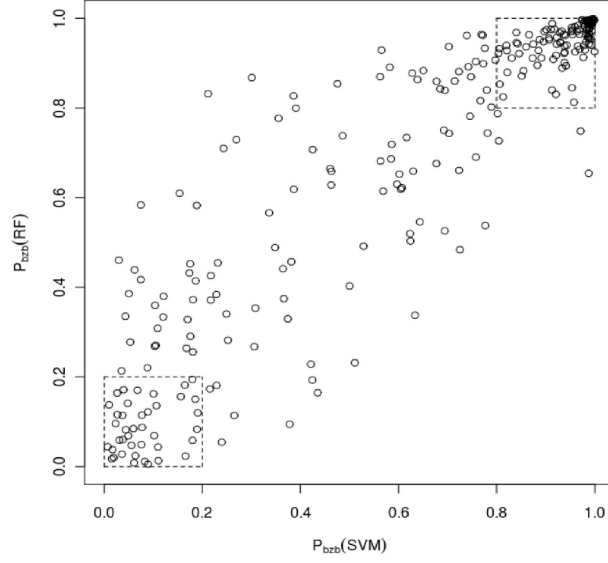


Figure 5.10: Decision threshold P_{bzb} obtained with RF versus P_{bzb} estimated by SVM for 257 AGU in the 2FGL. Dashed squares contain sources with common decision threshold over 0.8, classified with accuracy rates over 94%.

Table 5.4: Number of predicted AGU sources as a function of decision threshold.

	RF		SVMs		Both	
	bzb	bzq	bzb	bzq	bzb	bzq
$P > 0.5$	173	84	161	96	156	79
$P > 0.8$	129	46	112	63	106	39
$P > 0.95$	64	12	64	19	47	5

0.95). Individual algorithms and coincidences satisfying said conditions are included. Combining results from both classifiers and requiring $P > 0.5$, 235 (156 BL Lacs and 79 FSRQs) out of 257 objects are consistent with the properties of known γ -ray blazars. In order to place these results in context with identified/associated *Fermi* AGN, Fig. 5.11 shows the photon spectral index versus the flux ($E > 100$ MeV) of identified/associated BL Lacs and FSRQs overlaid with the AGU predictions from this work.

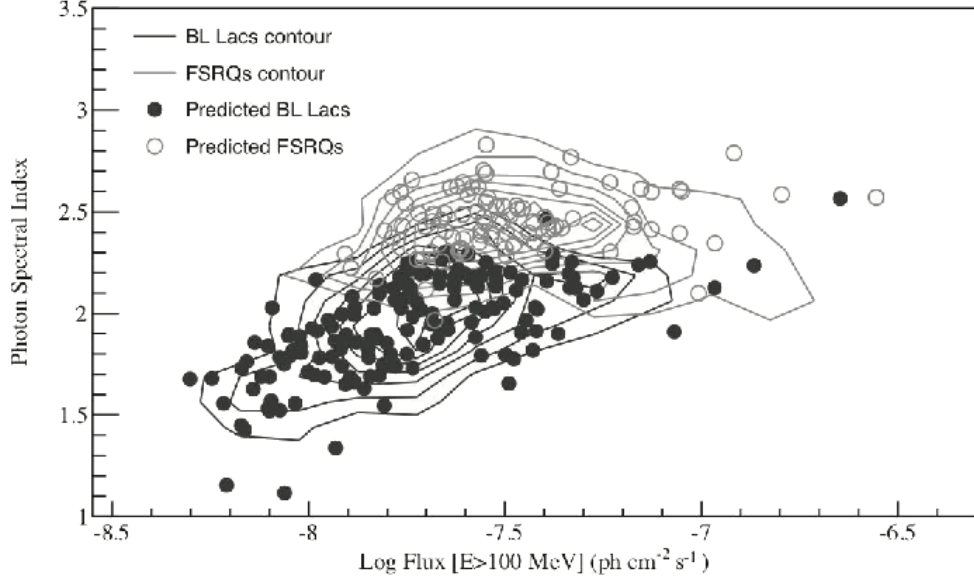


Figure 5.11: Photon spectral index versus log flux above 100 MeV for identified/associated BL Lacs (dark contour) and FSRQs (light contour). Predicted BL Lacs (filled circles) and predicted FSRQs (open circles) from the AGU dataset are shown over the contours.

5.3.2.3 Outlier detection and potential biases

Throughout, it is assumed that the classification of γ -ray AGN subclass falls along the two main blazar categories i.e. BL Lacs and FSRQs. Without final spectroscopy it is impossible to rule that other AGN subclasses are present in the AGU sample. As commented before, there is a minority of other subclasses in the 2FGL including Seyferts, radio galaxies and other AGN that have not been considered thus far. The main justification for ignoring further atomisation into subclasses is that blazars account for 97% of the identified/associated AGN sample. However, it is important to consider that a more complex mixture of AGN subclasses is possible. Fortunately, machine-learning algorithms excel at separating rare and unique objects from the dataset.

Adopting the method introduced in Sec. 5.3.1, a search is performed for AGU outliers that could potentially belong to other minority AGN subclasses. For this purpose, the outlying measure of each object is computed, defined as the reciprocal

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

sum of the squared proximities to all objects within its class. Outliers are defined as objects having small proximities to the rest of objects. Practically, RF returns proximities $prox(n, k)$ that represent the fraction of trees in which elements n and k fall in the same terminal node [170, 171]. Generally, anomalies are identified with outlier measures larger than 10. No source was found with such values, as a result there is no clear evidence of outliers in the AGU sample. For completeness, note that the highest values in the dataset correspond to 2FGL J1825.1–5231, 2FGL J1816.7–4942, and 2FGL J0022.3–5141 respectively.

This possibility is constraint further by retraining and testing the SVMs and RF algorithms with the full range of associated AGN subclasses present in the 2FGL. Given the size of the minority subclasses, care was taken to weight the classes appropriately to compensate the differences in the training sets. Taking into account additional AGN subclasses, at most 11 objects might belong to other AGN subclasses ($P < 0.6$). Therefore, there is no strong indication of contamination from additional subclasses. Taken together, both approaches limit the presence of other AGN subclasses in the AGU dataset. It is possible that the result simply reflects the small number statistics of additional AGN subclasses. A full characterization might improve in the future as *Fermi* expands its source catalogue.

5.3.2.4 Application to unassociated *Fermi* objects

In Sec. 5.3.1, class predictions were performed for the sample of unassociated *Fermi* sources at $|b| \geq 10^\circ$. In that initial approach, sources were sorted in broad AGN and pulsar categories. Given the success with further AGN subclasses, it may be interesting to extend the approach to all unassociated *Fermi* sources tagged as AGN. Using the same optimised models, the algorithms are applied to the 216 sources predicted as AGN in Sec. 5.3.1. The resulting predictions are shown in Table 5.5 with the same conditions adopted earlier. In this case, only 30% of the sources reach decision thresholds larger than $P > 0.8$ in both RF and SVM.

Table 5.5: Predictions for unassociated *Fermi* objects tagged as AGN by [203], ordered by RA.

Source	P_{bzb} (RF)	P_{bzb} (SVM)	Prediction
2FGL J0004.2+2208	0.15	0.11	bzq
2FGL J0014.3-0509	0.37	0.19	-
2FGL J0031.0+0724	0.97	0.94	bzb
2FGL J0032.7-5521	0.41	0.28	-
2FGL J0039.1+4331	0.87	0.99	bzb
2FGL J0048.8-6347	0.91	0.76	-
2FGL J0102.2+0943	0.90	0.89	bzb
2FGL J0103.8+1324	0.94	0.95	bzb
2FGL J0116.6-6153	0.97	0.99	bzb
2FGL J0124.6-2322	0.49	0.66	-
2FGL J0129.4+2618	0.19	0.05	bzq
2FGL J0133.4-4408	0.63	0.73	-
2FGL J0143.6-5844	1.00	0.99	bzb
2FGL J0158.4+0107	0.36	0.26	-
2FGL J0158.6+8558	0.06	0.07	bzq
2FGL J0200.4-4105	0.98	0.99	bzb
2FGL J0221.2+2516	0.99	0.99	bzb
2FGL J0226.1+0943	0.66	0.76	-
2FGL J0227.7+2249	0.89	0.95	bzb
2FGL J0239.5+1324	0.99	0.95	bzb
2FGL J0251.0+2557	0.37	0.19	-
2FGL J0305.0-1602	0.99	1.00	bzb
2FGL J0312.5-0914	0.93	0.69	-
2FGL J0312.8+2013	0.91	0.97	bzb

Note: The complete list of predictions is available at <http://www.gae.ucm.es/~thassan/agus.html>.

5. APPLICATIONS OF MACHINE LEARNING ALGORITHMS TO THE HIGH ENERGY ASTROPHYSICS

5.3.2.5 Discussion

RF and SVM classifiers were used to predict specific source subclasses for γ -ray AGN of uncertain type, by learning from features extracted from associated AGN in the 2FGL. Both algorithms are successful in capturing the properties of γ -ray AGN reaching accuracy rates of 85%. This effort shows that 235 out of 269 AGN of uncertain type have properties consistent with γ -ray BL Lacs and FSRQs, with decision thresholds over 0.8. Comparison of these predictions with the sample of associated AGN verify that similar populations are traced (Fig. 5.11). Nevertheless, without high-quality spectral observations, final counterpart association will have to wait for dedicated optical spectroscopy.

Apart from internal training and testing, results are cross-matched with a recent study showing that blazars can be recognised and separated from other extragalactic sources using infrared colours [202]. Class characterisation has been performed for *Fermi* AGN of uncertain type taking advantage of this *total strip parameter* traced by BL Lacs and FSRQs. The possibility of comparing obtained predictions with the source classes inferred from IR colours is ideal, as both methods are independent. For a subset of 54 overlapping sources listed in [202], predictions obtained in this work match in 85% of the objects with the $P > 0.5$ decision threshold, and the agreement rate improves to 93% for the 33 objects satisfying the $P > 0.8$ condition. The excellent agreement suggests that the method is viable and that infrared colours can not only recognise generic blazars but also provide information about specific blazar subclass i.e. BL Lac or FSRQ. More importantly, this cross-validation reinforces the power and possibilities of machine-learning algorithms as source classifiers in γ -ray astrophysics.

Even though the initial approach aimed to distinguish between BL Lacs and FSRQs, the possibility that other subclasses are represented within the AGU dataset was also considered. No clear outliers have been found within the latter. Training and testing after taking into consideration additional subclasses finds only 11 objects ($P < 0.6$) that might have been missed with a binary classification. This is consistent with findings indicating that additional AGN subclasses (Seyferts, radio galaxies and other AGN) account for a 3% of the whole AGN sample. There might be a bias introduced by the relative rarity of minority ob-

jects. Nevertheless, AGN of unknown type are most likely dominated by BL Lac or FSRQ, in agreement with [202].

The clear intent of this effort is to characterise the entire γ -ray population. It is expected that these results can help observers in future spectral and photometric endeavours aimed at classifying the entire AGN counterpart sample. Additionally, this work can help discriminate targets for follow-up studies of AGN at even higher γ -ray energies with ground-based IACTs (MAGIC, H.E.S.S., VERITAS). Viewing forward, γ -ray spectral features will be nicely complemented with the future CTA, expected to increase spectral coverage and sensitivity. The design of future survey pointing strategies for CTA [206] will also benefit from object lists such as the one presented in this work by boosting the AGN target pool available. In the shorter term, an obvious improvement that lies ahead is to incorporate multi-wavelength (radio, optical, X-ray) entries to complement individual classifying features.

Chapter 6

Conclusions

In this thesis, I have presented results on the sensitivity studies performed for the CTA collaboration, evaluating through the analysis of large-scale MC simulations the observatory performance, along with an estimation of its future potential on specific physics cases. Together with the testing and development of the analysis tools employed, these results are critical to understand CTA's future capabilities, the efficiency of different telescope placement approaches and the effect on performance of the construction site related to parameters such as the altitude or the GF. The Northern Hemisphere proposed construction sites were analyzed and evaluated, providing an accurate estimation of their capabilities to host the observatory.

As for the CTA layout candidates, an unbiased comparison of the different arrays proposed by the collaboration was performed, using the *Fermi*-LAT catalogs to forecast the performance of each array over specific scientific cases. In addition, the application of machine learning algorithms on γ -ray astronomy was studied, comparing alternative methods for energy reconstruction and background suppression and introducing new applications to these algorithms, such as the determination of γ -ray source types through the training of their spectral features.

The main conclusions of this work will be divided into the three main studied topics:

CTA sensitivity studies: The analysis presented here of both CTA-N and CTA-S candidates represents the most comprehensive study of CTA capabilities

6. CONCLUSIONS

performed by the collaboration to date. Experience gained with the improvement of this software will guide the future CTA analysis pipelines by comparing the attained sensitivity by alternative analysis chains. From these results, both CTA-N and CTA-S candidates “2N” and “2Q” fulfill the sensitivity, angular and energy resolution, effective area and off-axis performance requirements. MC simulations provide an useful test-bench for the different designs within the CTA project, and Prod-2 results demonstrate their correct implementation would attain the desired performance and potential scientific output.

The Schwarzschild-Couder (SC) IACT design, mounted on the SC-MST and SC-SSTs (both ASTRI and CHEC), show promising capabilities specially in terms of angular resolution and off-axis performance. Concerning the planned SC-MST extension, the halo approach wins over the interleaved option both in terms of sensitivity and angular resolution, providing specific constraints on the future CTA baseline. If the SC-MSTs extension is constructed in a layout with MSTs, the latter should be placed surrounding the center of the array, maximizing the amount of low energy events reconstructed by LSTs and MSTs while improving the direction reconstruction and collection area of the showers with energies within the CTA core energy range. Off-axis performance will be also boosted by a 20%, significantly improving CTA surveying capabilities and the chances of serendipitous discoveries.

Regarding the evaluation of the Northern Hemisphere site candidates presented in this work, the main performance differences were dominated by the strong Geomagnetic field affecting all proposed sites, overshadowing the effect of the construction altitude, which was initially thought to be dominant for the low energy performance. Higher elevations turned out to be double-edged, significantly decreasing high energy performance, which favoured layouts of moderate altitudes. The average geomagnetic field effect presented in this thesis was included in the CTA Site Evaluation Summary along with the calculated PPUTs. Given the small differences (within the calculated uncertainties) between the over-all sites capabilities, the main conclusion is that all studied sites fulfilled performance requirements.

Concerning the Spanish proposed sites, dedicated MC simulations have shown that moderate increased level of NSB, like the one observed in the Tenerife site, would have a negligible effect on performance above 100 GeV (with sensitivity losses below the 2%). In addition, moderate baseline modifications were applied

to the standard CTA-N layout to fit in the Observatorio del Roque de los Muchachos (La Palma, Spain). No sensitivity losses were reported except for a slight angular resolution loss above 10 TeV, expected to disappear with the use of more sophisticated direction reconstruction methods.

CTA forecast: A software package was developed by the author of this work to calculate the significance level of detection attained by CTA observations on *Fermi*-LAT sources using the IRFs of different layout candidates. The main conclusions regarding the physics cases studied in this work are listed below:

- Source populations: Considering the galactic and extragalactic populations accesible to the future CTA, layouts with higher number of LSTs, such as “B” or “E” arrays, would detect a larger amount ($\sim 10\%$) of sources with equivalent observation times. Sites located at moderate altitudes (such as the Namibian site, at 1640 m) are also preferred, improving coverage by a 5-10%. In any case, all considered layouts achieve outstanding results, with galactic populations of more than 70 sources with individual exposure times of 5 hr. Under the consideration of an extragalactic survey, 20 sources would be detectable with 0.5 hr of individual observation, increased to 80 objects if 5 hr are considered. This forecast was included in [206].
- AGNs: By performing follow-up observations of known *Fermi* sources at lower frequencies, more than 200 AGNs could be accessible to CTA with exposure times lower than 50 hours, irrespective to observability conditions. Blazars at distances up to $z \approx 2$ are expected to be detected in their quiescent state, although more distant sources are expected to be detected during flaring states. The CTA-N site will greatly affect these numbers, as more than the 50% of these sources are only observable from the Northern Hemisphere. Favoured layouts are again “B” or “E” arrays, built at moderate altitudes. These results were part of [207].
- Pulsars: Under the hypothesis that the existence of VHE-tails is a universal feature in γ -ray pulsars, ≈ 20 pulsars would be detectable by CTA. This represents a large fraction (up to 40%) of the brightest Fermi pulsars. CTA

6. CONCLUSIONS

configurations with more LSTs are preferred, as pulsar detection falls in the low energy range (30 to 200 GeV). High construction altitudes would allow the detection of more pulsars if their emission was consistent with an exponential cut-off, while lower site altitudes would increase the detected VHE γ -ray pulsars population if they had similar behaviour as the Crab pulsar in the VHE range. No γ -ray emission was considered from any PWN, so these results may be too optimistic (although valid for layout comparison). These were included in [208].

- DM prospects: The Galactic Center will be the best candidate for constraining DM models considering γ -ray emission through WIMPs annihilation, although a detection in this regime could be attributed to other emitters. As for dSph, Segue 1 will be the most encouraging object for this kind of study and it is only reachable from the Northern hemisphere. Predictions seem to show that boost factors $B_F \approx 15$ are required for a detection for observations of 250 hr considering a 300 GeV WIMP purely annihilating into $\tau^+\tau^-$. For WIMP masses below 1 TeV, “B” or “E” arrays are preferred while other layouts with greater VHE capabilities (such as “C”) are favored for higher masses. In any case, CTA will be the only facility capable of providing competitive constraints at these energies, with the possibility of an unequivocal DM detection. These results were published in [23].

Machine learning applications: In this work, machine learning algorithms performance and a wide range of applications were studied, from high-performance event classification for Cherenkov telescopes to novel applications of source classification of γ -ray source catalogs, which will be crucial in the CTA era.

Improvements in the event classification will be critical in order to detect the faintest sources, specially improving the low and core energy ranges. Energy estimation is proofed to be boosted in the whole range of energies by using machine learning algorithms by a factor 2. Concerning γ -hadron separation, several algorithms were tested: the purely randomized Random Forest resulted to outperform considered alternatives, such as Neural Networks or Support Vector Machines. Note that electrons are the main source of background for energies between 100

GeV and 10 TeV, so improved γ - e^- separation methods, with the inclusion of new parameters to discern between them, could be the key to improve the future CTA sensitivity.

With the growing *Fermi*-LAT catalogs and the future addition of HAWC and CTA detections, source type classification algorithms will be a necessity once a new much expanded very high energy source catalog is produced. In this work, two classification methods are proposed showing excellent results. Unassociated sources with positions coincident with the galactic plane are classified into AGNs or pulsars with accuracy rates of 97.1%. No clear outliers were found, showing all sources could be well described by these two classes. As a second step, further classification was performed to understand all the intricacies of the AGN population, discerning between FSRQ and BL Lac sub-types. In this case, average accuracy rates of 85% are obtained, given the similarities of both AGN classes. The results of both studies were published in the Monthly Notices of the Royal Astronomical Society (MNRAS) [203, 209].

6. CONCLUSIONS

Appendix A

Extended Atmospheric Showers

As already mentioned in chapter 2, our atmosphere is constantly being hit by high energy cosmic particles. This collisions produce cascades of sub-particles called Extended Air Shower (EAS). This appendix should be considered as an introduction to the physics involved in EAS, to help the reader understand where Cherenkov photons are coming from, and the differences between hadronic, leptonic and electromagnetic showers. These differences are of great importance for background suppression in the IACT technique.

First, in Sec. A.1 a short explanation of the Cherenkov radiation is provided, to understand EAS photons origin. Then, EAS particle production along the atmosphere is described for γ -ray and Cosmic Ray (CR) induced showers. This appendix ends with a brief explanation to the Geomagnetic field effect over pure electromagnetic cascades.

Further detailed information about extended air showers can be found in [210].

A.1 The Cherenkov radiation

Cherenkov radiation is the electromagnetic radiation a charged particle emits while traveling in a dielectric medium faster than the velocity of light. It was discovered and named after the Soviet scientist Pavel Alekseyevich Cherenkov, Nobel Prize

A. Extended atmospheric showers

winner in 1958 due to the experimental detection of this effect.

The origin of Cherenkov radiation is shown in Fig. A.1. Along the path of the charged particle, the medium becomes electrically polarized. If in a certain medium of refraction index n the particle travels at inferior velocity than the local velocity of light $v < c/n$, the disturbed medium relaxes back to an equilibrium as the particle transits through, as the net polarization field is null due to the symmetric arrangement of the dipoles. In the case where the particle velocity is greater than $v > c/n$, the medium is not able to relax back elastically and the polarization gets asymmetric (as shown in Fig. A.1a), leaving a disturbance in the wake of the particle radiating as a coherent shock-wave (as shown in Fig. A.1b) the energy contained in the disturbance, reorienting the asymmetric dipoles, emitting Cherenkov photons.

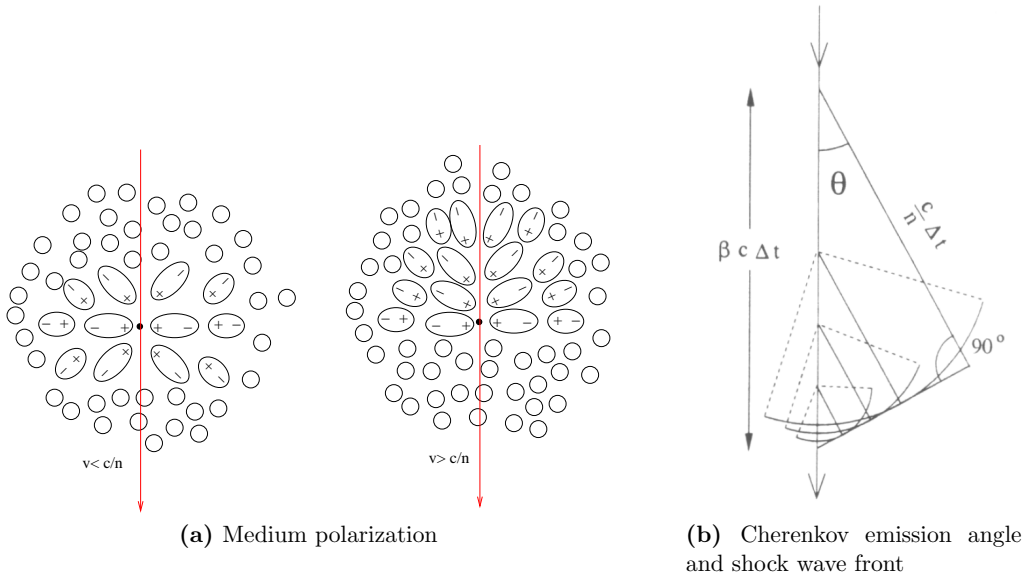


Figure A.1: Fig. A.1a: Medium polarization due to a charged particle traveling slower (*left*) and faster (*right*) than the local phase velocity of light. Fig. A.1b: Shock wave of in-phase reorientation of dipoles, origin of the Cherenkov photons. Figures extracted from [211].

As depicted in Fig. A.1b, Cherenkov radiation is emitted with a certain angle θ with respect to the charged particle direction named *Cherenkov angle*. Applying the Huygens principle, and considering the charged particle recoil due

to the photon emission, the *Cherenkov angle* can be expressed as:

$$\cos\theta = \frac{1}{n\beta} + \frac{\hbar k}{2p} \left(1 - \frac{1}{n^2}\right), \quad (\text{A.1})$$

where β is $\beta = v/c$ and k and p are the Cherenkov photon and charged particle momenta respectively. As the photon momentum can be neglected compared with the charged particle momentum, Eq A.1 can be approximated by the generally used expression of the *Cherenkov angle*:

$$\cos\theta = \frac{1}{n\beta}. \quad (\text{A.2})$$

From Eq. A.2 we deduce that θ increases with the velocity of the charged particle which has to exceed $\beta_{min} = 1/n$ for the radiation to be emitted. Another consequence of this expression is that the maximum angle, corresponding to ultra-relativistic particles with $\beta_{max} = 1$, is $\theta_{max} = \cos^{-1}(n^{-1})$. The threshold energy a particle needs in order to emit Cherenkov photons is:

$$E_C^{thr} = \frac{m_0 c^2}{\sqrt{1 - \beta_{min}^2}} = \frac{m_0 c^2}{\sqrt{1 - (1/n)^2}}, \quad (\text{A.3})$$

where m_0 is the rest mass of the particle. When highly relativistic particles travel down the atmosphere we need to take into account that the refraction index varies with altitude. In a first approximation this dependence can be expressed by:

$$n(h) = 1 + n_0 e^{-\frac{h}{h_0}}, \quad (\text{A.4})$$

considering an isothermal atmosphere where $n_0 = 2.9 \times 10^{-4}$. Using both equations we obtain the threshold energy for Cherenkov light emission as a function of the altitude:

$$E_C^{thr} \approx \frac{m_0 c^2}{\sqrt{2n_0}} e^{\frac{h}{2h_0}}. \quad (\text{A.5})$$

A. Extended atmospheric showers

Using this equation for particles usually found in EAS, the threshold energy at sea level for electrons, muons, pions and protons is 21 MeV, 4.4 GeV, 5.8 GeV and 39 GeV respectively. Eq. A.2 turns into:

$$\theta_{max} \approx \sqrt{2n_0} e^{-\frac{h}{2h_0}}, \quad (\text{A.6})$$

That is the maximum Cherenkov angle, shown in Fig. A.4a, and can be considered the regular emission angle along the EAS development. For high altitude it is less than 0.5° , gradually broadening at lower altitudes up to 1° at 5 Km and reaching 1.4° at sea level.

Regarding the Cherenkov spectrum, the photons generated by a particle with charge number z & velocity β can be expressed as [212]:

$$\frac{d^2N}{dx d\lambda} = \frac{2\pi\alpha z^2}{\lambda^2} \left(1 - \frac{1}{\beta^2 n^2(\lambda)} \right). \quad (\text{A.7})$$

In the case of Cherenkov emission in the atmosphere, introducing Eq. A.4 in the previous equation, shows the Cherenkov spectrum ranges between the infrared and ultraviolet. Part of this emission is absorbed (by ozone, H_2O and CO_2) and scattered by air molecules and aerosols, leading to a final spectrum observed in the ground with its maximum located at $\lambda \approx 330$ nm (light blue).

A.2 Extended atmospheric showers

The atmosphere is constantly absorbing the impact of highly energetic cosmic particles, giving a crucial protection for the existence of life on Earth. Particles with enough energy, produce cascades of secondary particles called *Extended Atmospheric Showers* (EAS). These secondary particles travel faster than the local speed of light, emitting Cherenkov radiation, as explained in Sec. A.1.

All EAS follow certain general geometrical properties: they typically measure several kilometers in length and some hundreds of meters in width. The height where the maximum number of particles is generated is usually between 8 and 12 km in elevation, corresponding to atmospheric mass fraction between 0.2 and 0.3. Depending on the characteristics of the primary particle EAS develop in different

ways, giving rise to different patterns of Cherenkov radiation. The two main types of EAS are electromagnetic, generated from incoming γ -rays, and hadronic, produced by impinging cosmic rays. These differences are crucial for the IACT technique, as the signal is notably dominated by background, and its rejection relies on the difference between electromagnetic and hadronic induced EAS.

A.2.1 Gamma-ray induced extended atmospheric showers

When incoming γ -rays interact with the atmosphere pure electromagnetic showers are generated, converting an atmospheric nucleus into an electron-positron pair. These first generation of particles also interact with atmospheric nuclei producing high energy bremsstrahlung photons decreasing their energy. These bremsstrahlung photons are energetic enough to continue producing electron-positron pairs, so this process is repeated cyclically producing a cascade of electron, positrons and high energy photons (see Fig. A.2). Although photons are able to produce electron-positron pairs as long as they have energy exceeding twice the rest energy ($m_e c^2$) of an electron ($2 \times 0.511 = 1.022 \text{ MeV}$), below a certain *critical energy* $E_c \simeq 100 \text{ MeV}$ the ionization and Compton scattering of the generated electrons and positrons become dominant as energy decreases, reducing bremsstrahlung emission and limiting high energy photons production.

To understand the characteristics of EAS we need first to define certain parameters: The *radiation length* ξ_0 measures the mean distance in which the energy of the charged particle is reduced by a factor e . In a similar fashion, the *mean free path* of photons is defined as the mean distance covered by a population after which the total number of photons is reduced by a factor e due to electron-positron pair production.

The development of an electromagnetic shower can be qualitatively understood using a toy model due to Heitler [213]. The shower is considered to develop in discrete steps, of length equal to one radiation length. In the case of ultra-relativistic electrons this length is considered equal to the mean free path of photons of equal energy, which is actually a good approximation, since $\xi_{\text{brems}} \sim 7/9 \xi_{\text{pp}}$. Also, the distribution of energy between the charged particle and the emitted bremsstrahlung photon is considered symmetric. In this way the number of surviving electrons,

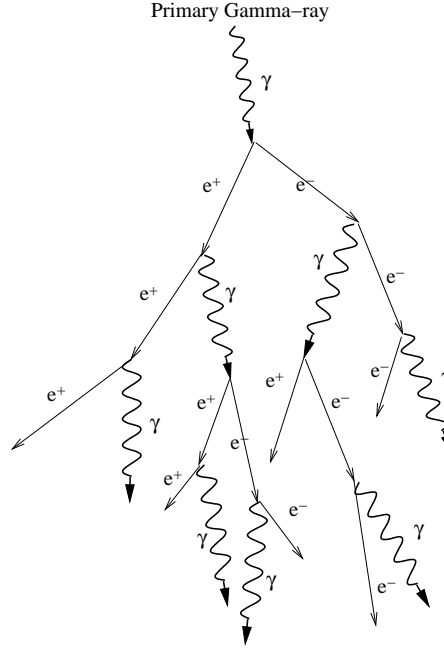


Figure A.2: Development of a pure electromagnetic cascade, generated from an impinging primary γ -ray, extracted from [10].

positrons and photons of the n -generation of radiation-absorption is 2^n , and their mean energy can be expressed, in terms of the initial primary particle energy E_0 , as $E_0/2^n$. In the shower maximum X_{max} the number of surviving electron, positron and photons is approximately E_0/E_c , where E_c is the previously defined critical energy so it can be approximated to:

$$X_{max} = \xi_0 \ln \frac{E_0}{E_c}, \quad (\text{A.8})$$

The development of electromagnetic showers in the atmosphere was addressed by Rossi and Greisen (see [214] for further details), resulting in the simplified *Greisen* formula, where the number of electrons above the critical energy E_c can be expressed as:

$$N_e(s) = \frac{0.31}{\sqrt{\ln(E_0/E_c)}} e^{T(-1.5 \ln(s))}, \quad (\text{A.9})$$

where s is the *shower age*, and T is the *atmospheric depth* expressed in radiation

A.2. Extended atmospheric showers

lengths. The shower age s is a function of the atmospheric depth T , and indicates the degree of development of the shower having a value of 0, 1 and 2 in the first interaction, shower maximum and extinction point respectively. It is defined as:

$$s = \frac{3T}{T + 2\ln\left(\frac{E_0}{E_c}\right)} \quad (\text{A.10})$$

Generally $N_e(T)$ is called the longitudinal development of the shower, shown in Fig. A.3 for different E_0/E_c values. The lateral distribution $\rho_e(r)$ is defined as the electron-positron density as a function of r , the distance to the shower axis, for a given elevation. Using the Nishimura-Kamata-Greisen formula [215] can be expressed as:

$$\rho_e(r) = \frac{N_e}{r_M^2} \left(\frac{r}{r_M}\right)^{s-2} \left(1 + \frac{r}{r_M}\right)^{s-4.5} \frac{\Gamma(4.5-s)}{2\pi\Gamma(s)\Gamma(4.5-2s)} \quad (\text{A.11})$$

where the distance in the shower axis r is expressed in units of the Molière radius $r_M = 79 \text{ m}$ at sea level, and Γ is the Gamma function ($\Gamma(n) = (n-1)!$).

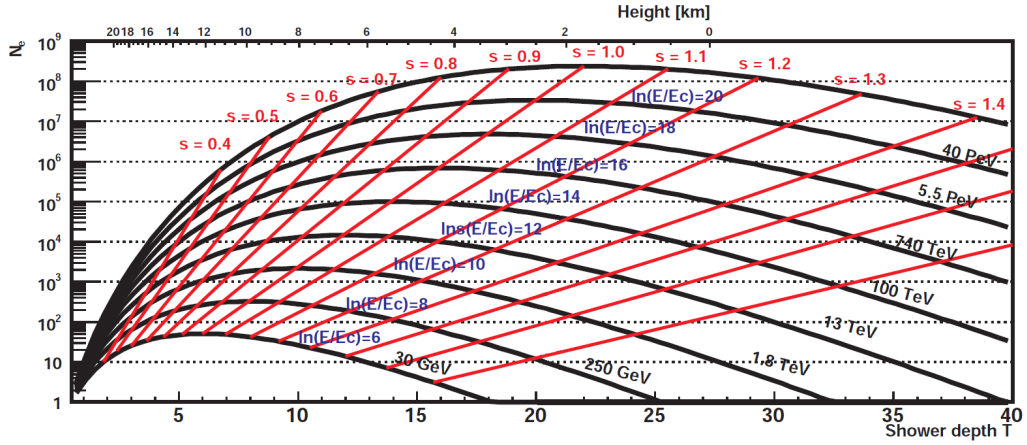


Figure A.3: γ -ray induced EAS longitudinal development under Greisen approximation, described by the shower size N_e over the shower depth T . Black lines correspond to different primary γ -ray energy, while red lines define points of equal shower age s . Figure courtesy of [9].

In order to account for the Cherenkov emission generated in the EAS, lateral distribution of emitted photons can be computed adding the contribution of every

A. Extended atmospheric showers

generated super-luminic particle. In the case of particles traveling through the atmosphere perpendicular to the ground, the distance between the trajectory of the particle and produced photons can be expressed as:

$$R_C(h) = (h - h_{obs}) \tan \theta_{max}, \quad (\text{A.12})$$

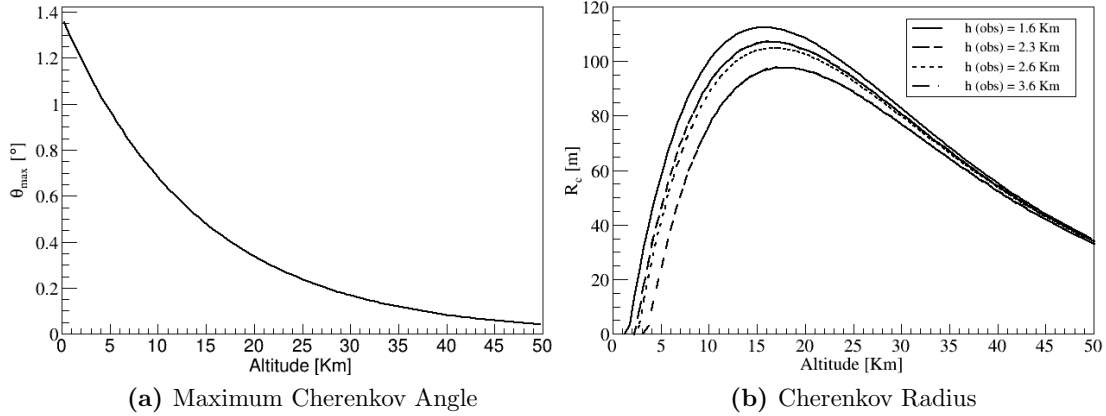


Figure A.4: Cherenkov Angle (Fig. A.4a) and distance R_c (Fig. A.4a) of a perpendicularly impinging γ -ray along the atmosphere. From Fig. A.4b several altitudes are shown, corresponding to different simulated altitudes by the CTA Monte Carlo group, with maximum R_c values ranging between $R_c(3.6Km) \approx 95m$ and $R_c(1.6Km) \approx 110m$

where h_{obs} is the observation altitude and θ_{max} the maximum Cherenkov angle of eq. A.6. The small angle Cherenkov photons are emitted at high altitudes, fix a relatively constant light pool size for pure electromagnetic cascades on the ground, where the maximum density of photons is located at a distance R_c of the shower core. Depending on the elevation, this radius ranges between $\approx 90m$ at very high altitudes (5000 m) and $\approx 120m$ at sea level.

A.2.2 Hadron induced extended atmospheric showers

Gamma rays are not the only high energy particles colliding with the atmosphere. In fact, γ -ray induced EAS correspond to less than the 0.01% of the total generated showers. The vast majority of EAS are originated by Cosmic rays. They develop from the collision of high energy nuclei, mostly protons, with an atmospheric nu-

cleus, generally N or O, and differ significantly from pure electromagnetic cascades (see Fig. A.5). In the first interaction, fragments of the original nuclei and many kinds of secondary particles are generated, mainly pions (π^0, π^+, π^-), and in a lower degree kaons (K^+, K^-) and nucleons. The lateral dispersion of the cascade is dominated by the transverse momentum of the secondary particles created in the collision, unlike pure electromagnetic cascades, dominated by Coulomb dispersion, which does not create transverse momentum. Although protons are the main component of the cosmic rays colliding with the atmosphere, a wide range of particles and nuclei have been observed to contribute in the Cosmic Ray Spectrum, shown in Fig. A.6.

Hadronic cascades have three main components. The electromagnetic component, the *hadronic core*, and the muon and neutrino component.

Most of particles generated in hadronic cascades, around 90%, are pions and one out of 3 is a π^0 , which decay in 2 γ -rays creating pure electromagnetic cascades. These are the main contributors to the EM component, absorbing $\sim 30\%$ of the collision energy.

The hadronic core is formed by high energy nucleons and charged mesons. Generated nucleons keep colliding with atmospheric nuclei in a similar fashion as the primary collision as long as they have more energy than the pion production threshold (~ 1 GeV). Charged mesons collide with atmospheric nuclei creating new sub-particles, or desintegrate into muons and neutrinos ($K^\pm \rightarrow \pi^\pm + \pi^0$, $\pi^\pm \rightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$).

The last component are the muons and neutrinos generated from hadronic collisions. Muons loose energy through ionization and Cherenkov emission, and can also desintegrate into electrons and neutrinos ($\mu^\pm \rightarrow e^\pm + \nu_e(\bar{\nu}_e) + \bar{\nu}_\mu(\nu_\mu)$) although their high Lorentz factor together with their relatively large mean lifetime, produce time dilation, causing many muons to reach ground level before they disintegrate.

In a similar fashion as in electromagnetic EAS, a simple model can be used to approximate the behaviour of hadron induced cascades. This model assumes that a cascade induced by a nucleus of mass A and energy E_0 can be considered equivalent to A nuclei of energy E_0/A interacting with the atmosphere independently. Using this approximations, similar to eq. A.8, the shower maximum of hadron induced

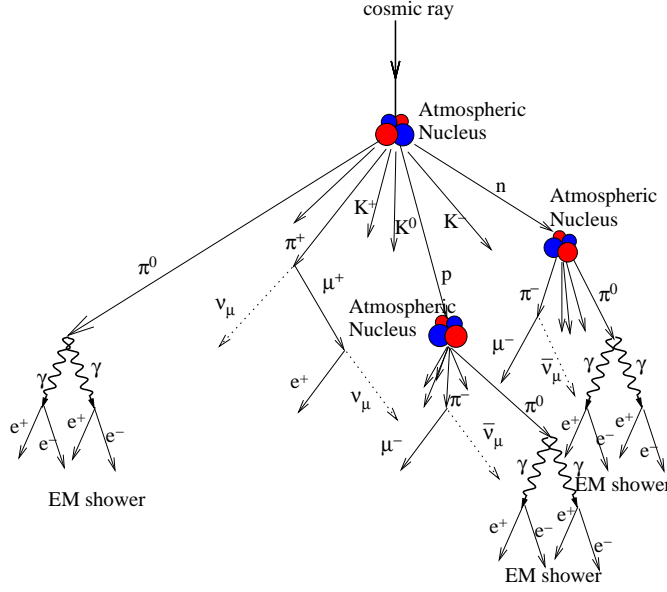


Figure A.5: Schematic figure of an hadronic shower. Since heavy hadrons can present high transverse momentum that is transferred to the products, the shower width is broadened as compared to a pure electromagnetic EAS. Figure extracted from [9].

cascades can be:

$$X_{max} \propto \xi_N \ln \left(\frac{E_0}{AE_c} \right). \quad (\text{A.13})$$

Where E_c is the critical energy and ξ_N corresponds to the hadron interaction length, defined as the mean path required to reduce the number of relativistic charged particles by the factor $1/e$ as they pass through matter ($\xi_N \simeq 80g/cm^2$ for a 1 TeV proton). Since ξ_N is larger than ξ_0 , we deduce the altitude of the first interaction is higher compared with pure electromagnetic cascades. From eq. A.13 we infer that cascades generated by primary particles of equal energy develop at higher altitudes if they are more massive.

Unlike γ -ray cascades, hadron induced EAS are more complex and Monte Carlo simulations need to be performed. These simulations use the results of the current generation of hadron colliders, which are constrained by their energy range and their limited sensitivity for collisions with low transverse momentum. For the first interactions in very energetic hadronic cascades, extrapolations need to be assumed adding certain degree of uncertainty.

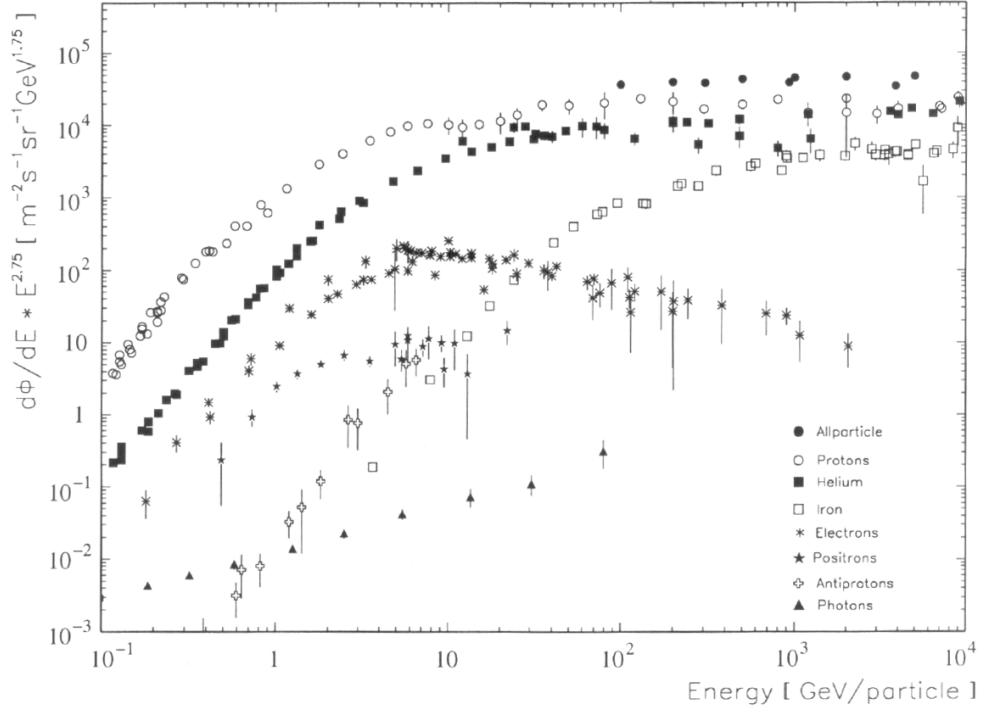


Figure A.6: Different contributions to the Cosmic Ray spectrum between 100 MeV and 10 TeV collected by [216]. Figure extracted from [211].

A.2.3 Electron induced extended atmospheric showers

Within the cosmic ray radiation there are non-hadronic components that also generate EAS. These particles, whose spectra upon incidence on the atmosphere are shown in Fig. A.6, are high energy electrons and positrons which in collision with the atmosphere generate pure electromagnetic cascades in a process similar to the one of γ -ray initiated EAS. In comparison with the hadronic component, electrons and positrons correspond to less than a 1% of the total cosmic ray flux, but they become an important source of background for atmospheric Cherenkov detectors due to the similarities they share with γ -rays.

Cascades are generated when an impinging high energy electron or positron interacts with an atmospheric nucleus producing a high energy bremsstrahlung photon, and creating an electromagnetic cascade, already explained in section A.2.1. Although γ -ray and electron-positron induced showers share their nature and the same processes take place in their development, there are certain differences among

A. Extended atmospheric showers

them coming from the differences between primary particles.

In case of e^\pm induced cascades, the primary particles are charged, so as soon as they enter the atmosphere Cherenkov emission starts. The primary particle also loses energy through bremsstrahlung at higher altitudes, as shown in Fig. A.7, developing the cascade earlier and reaching the shower maximum at higher altitudes (and lower atmospheric depths).

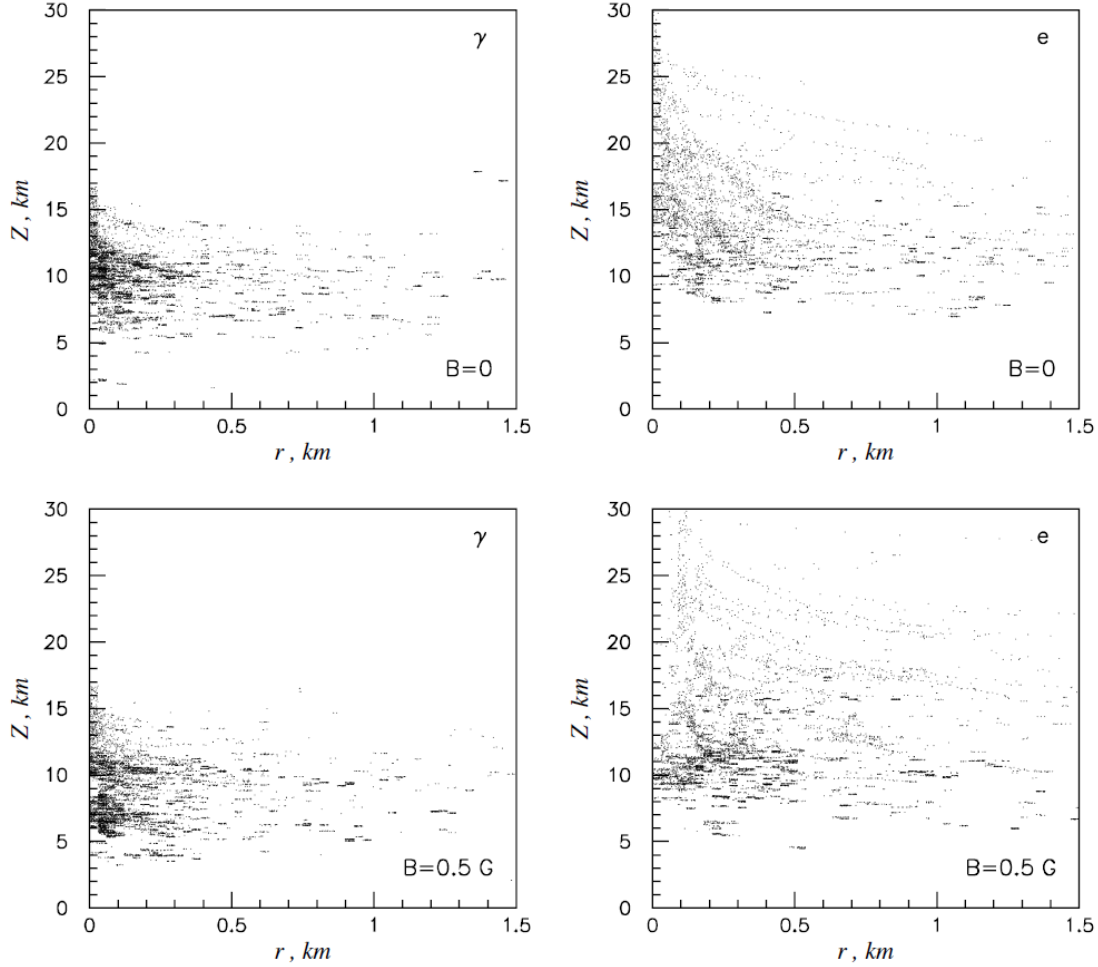


Figure A.7: Location of generated secondaries (height versus distance from the shower core) of a primary particle of 10 GeV by [217]. Gamma induced showers correspond to the left side plots while electron induced showers to the right side. No geo-magnetic effect is considered in the upper side plots while $B = 0.5 \text{ G}$ is imposed in the lower side

Comparing the depths of the maximum between γ and e^- induced showers

with energy E , done by [217]:

$$X_{max}^{\gamma(e)}(E) = 1.01\xi_0 \left[\ln \frac{E}{E_c} - n \right], \quad (\text{A.14})$$

where X_{max} is the depth of the shower maximum height of γ and e^- induced showers when $n = 0.5$ and 1 respectively, ξ_0 the radiation length and E_c the critical energy defined in Sec. A.2.1. This equation shows there is a constant difference between the depths of their shower maximum of $\Delta(E) = 0.505X_0 \approx 37.1gcm^{-2}$. This $\Delta(E)$ corresponds to a difference in altitude that varies with the height the shower maximum occurs which, in turn, depends on the energy of the primary particle. Higher differences between γ and e^- showers appear at low energy showers, corresponding to higher maximum heights (see Fig. A.3).

If we consider the atmosphere density decreases exponentially with height, we can express $\Delta(E)$ like:

$$\Delta(E) = h_s \ln \left[1 + \frac{1}{2(\ln \frac{E}{E_c} - 1)} \right] \simeq \frac{h_s}{2(\ln \frac{E}{E_c} - 1)}, \quad (\text{A.15})$$

where h_s is the scale-height of the atmosphere ($h_s \simeq 7.1km$). Eq. A.15 shows $\Delta(E)$ decreases logarithmically with the energy, so differences between γ and e^- induced showers will be harder to measure as energy increases.

Another difference in the development of e^- induced EAS is the deviation suffered by charged particles from the geo-magnetic field. This effect, more significant for e^- induced EAS, will be analyzed further in section A.3.

A.3 Geomagnetic field effect on *EAS*

Along the development of atmospheric showers, Earth's Geomagnetic Field (GF) exerts Lorentz forces on the generated charged particles, bending the trajectories. The effect is stronger on low energy particles and affect mostly the e^-/e^+ pairs generated in the EAS. This forces generates a broadening effect on the lateral development of air showers not negligible compared with Coulomb scattering[218].

This force affects every generated charged particle and depends on the zenithal

A. Extended atmospheric showers

(θ) and azimuthal (ϕ) of the shower axis and the local magnetic field intensity.

The magnetic force affects all charged particles in the shower and depends on the angle formed between their trajectories and the GF. Since in first approximation, the direction of the trajectories can be approximated to the shower axis and the force is perpendicular to both the GF and the axis, we have that:

$$\vec{F} \propto \vec{B}_\perp = B_z \sin \theta \sin \phi \vec{i} + (B_z \cos \theta - B_x \sin \theta \cos \phi) \vec{j} + B_x \sin \theta \sin \phi \vec{k} \quad (\text{A.16})$$

where θ and ϕ are respectively the zenithal and azimuthal angle of the shower axis and \vec{B} is the local GF intensity, fixed by the location of the observatory. Eq. A.16 shows a direct dependency between the Lorentz effect on e^\pm and the direction of the particles, disrupting the azimuthal symmetry of pure electromagnetic cascades and creating East-West anisotropies in the footprint. This effect is shown in Fig. A.8, using the GF intensity of La Palma ($B_x = 30.2 \mu T$, $B_z = 24.2 \mu T$) and calculating the average footprint of 100 γ -rayinduced showers of 50 GeV impinging vertically ($\theta = 0$) at an altitude of 2.2 km.

The Geomagnetic field effect on extended air showers becomes more prominent the longer charged particles travel through the atmosphere. There are two cases where this occurs:

- *High zenith angle:* Particles impinging the atmosphere with high zenith angles (high angles with respect to the ground) travel longer distances in the atmosphere after the first interaction, making the spread of generated e^\pm more significant, leading to an increased dissemination of Cherenkov photons at ground level. As the zenith angle increases, the density of Cherenkov photons for low radial distances decreases proportionally to the orthogonal GF intensity, as can be observed in Fig. A.9.
- *Electron induced cascades:* In the case of e^-/e^+ induced EAS, as it was introduced in Sec. A.2.3, showers develop earlier, as the altitude of the first interaction is significantly higher (see Eq. A.14). This effect, although not very significant, creates higher anisotropies in the east-west direction of the footprint of e^-/e^+ induced cascades with respect to γ -rays. In addition, while

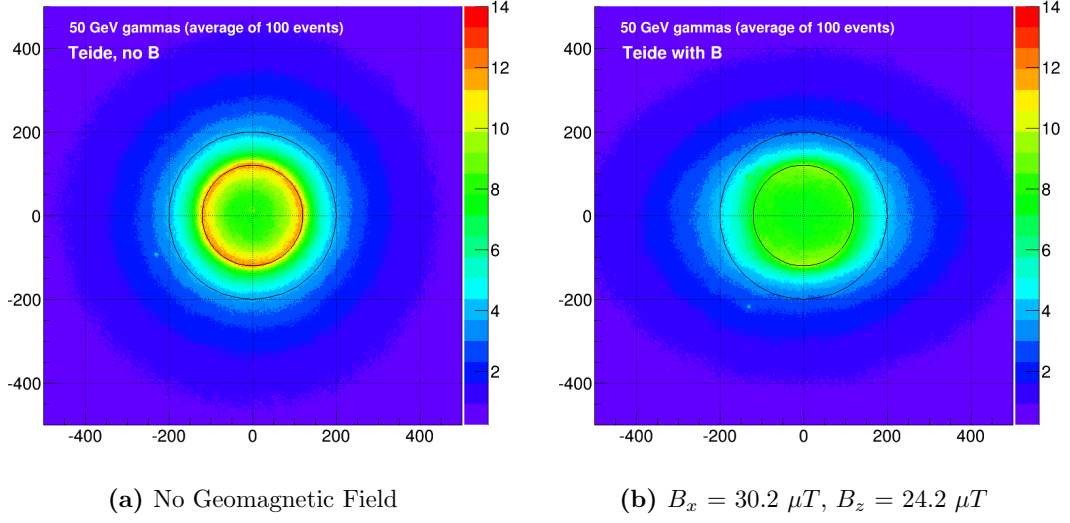


Figure A.8: Geomagnetic effect on the distribution of Cherenkov photons at ground level emitted by a γ -ray induced showers of $E = 50$ GeV with $\theta = 0$. Note that in La Palma South impinging γ -rays develop along the magnetic field lines, while North γ -rays 90° perpendicular to it, maximizing the effect over generated e^-e^+ pairs. From private communication with A. Moralejo.

emitting Cherenkov photons, impinging electrons bend their trajectories due to the geo-magnetic field at very high altitudes while γ -rays are not affected. Both effects can be observed in Fig. A.7.

For a detailed analysis of the Geomagnetic field effect on EAS and IACT observations, see [219].

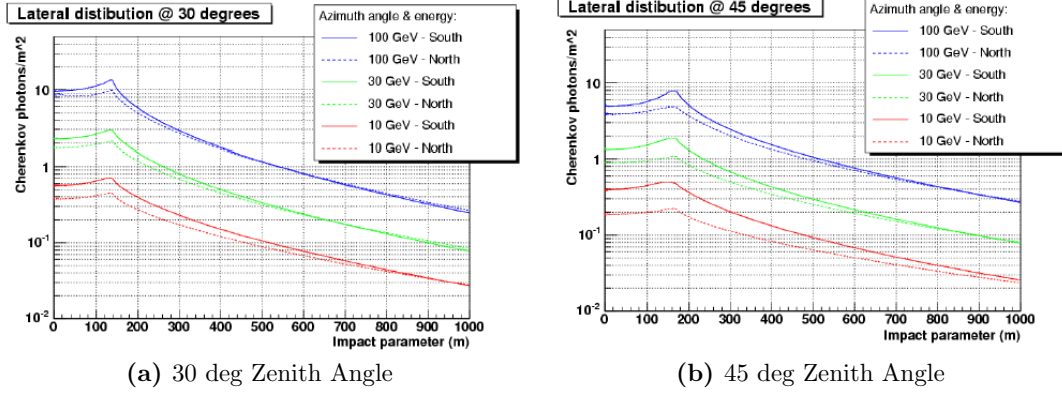


Figure A.9: Lateral distribution of Cherenkov photons at ground level of γ -ray induced showers of South ($\theta=0$) and North ($\phi=180^\circ$) directions. Showers were generated using the MAGIC telescopes location, with $B_x = 30.2 \mu T$, $B_z = 24.2 \mu T$. Figure extracted from [219].

Appendix B

Sensitivity studies for the Cherenkov Telescope Array

Since the creation of the first telescope in the 17th century, every major discovery in astrophysics has been the direct consequence of the development of novel observation techniques, opening new windows in the electromagnetic spectrum. After Karl Jansky discovered serendipitously the first radio source in 1933, Grote Reber built the first parabolic radio telescope in his backyard, planting the seed of a whole new field in astronomy. Similarly, new technologies in the 1950s allowed the establishment of other fields, such as the infrared, ultraviolet or the X-rays.

The highest energy end of the electromagnetic spectrum, the γ -ray range, represents the last unexplored window for astronomers and should reveal the most extreme phenomena that take place in the Universe. Given the technical complexity of γ -ray detection and the extremely relative low fluxes, γ -ray astronomy has undergone a slower development compared to other wavelengths. Nowadays, the great success of consecutive space missions together with the development and refinement of new detection techniques from the ground, has allowed outstanding scientific results and has brought gamma-ray astronomy to a worthy level in par with other astronomy fields.

This work is devoted to the study and improvement of the future Cherenkov

B. SENSITIVITY STUDIES FOR THE CHERENKOV TELESCOPE ARRAY

Telescope Array (CTA), the next generation of ground based γ -ray detectors, designed to observe photons with the highest energies ever observed from cosmic sources. These results on the sensitivity studies performed for the CTA collaboration evaluate the observatory performance through the analysis of large-scale Monte Carlo (MC) simulations, along with an estimation of its future potential on specific physics cases. Together with the testing and development of the analysis tools employed, these results are critical to understand CTA's future capabilities, the efficiency of different telescope placement approaches and the effect on performance of the construction site, related to parameters such as the altitude or the Geomagnetic Field (GF). The Northern Hemisphere proposed construction sites were analyzed and evaluated, providing an accurate estimation of their capabilities to host the observatory.

As for the CTA layout candidates, an unbiased comparison of the different arrays proposed by the collaboration was performed, using *Fermi*-LAT catalogs to forecast the performance of each array over specific scientific cases. In addition, the application of machine learning algorithms on γ -ray astronomy was studied, comparing alternative methods for energy reconstruction and background suppression and introducing new applications to these algorithms, such as the determination of γ -ray source types through the training of their spectral features.

The analysis presented here of both CTA-N and CTA-S candidates represents the most comprehensive study of CTA capabilities performed by the collaboration to date. Experience gained with the improvement of this software will guide the future CTA analysis pipelines by comparing the attained sensitivity by alternative analysis chains. From these results, both CTA-N and CTA-S candidates “2N” and “2Q” fulfill the sensitivity, angular and energy resolution, effective area and off-axis performance requirements. MC simulations provide an useful test-bench for the different designs within the CTA project, and Prod-2 results demonstrate their correct implementation would attain the desired performance and potential scientific output.

Appendix C

Estudios de sensibilidad para el Cherenkov Telescope Array

Desde la creación del primer telescopio en el siglo XVII, cada gran descubrimiento en la astrofísica ha sido la consecuencia directa del desarrollo de nuevas técnicas de observación, abriendo nuevas ventanas en el espectro electromagnético. Después de que Karl Jansky descubriera por casualidad la primera fuente radio en 1933, Grote Reber construyó el primer telescopio parabólico en el patio trasero de su casa, plantando la semilla de un nuevo campo en la astronomía. Del mismo modo, las nuevas tecnologías desarrolladas en la década de 1950 permitieron el establecimiento de otros campos, como el del infrarojo, ultravioleta y rayos-X.

Los fotones más energéticos del espectro electromagnético, los rayos γ , representan la última ventana inexplorada por los astrónomos, revelando los fenómenos más extremos que tienen lugar en el Universo. La complejidad técnica que supone la detección de rayos γ y los flujos tan bajos a los que se emiten, ha supuesto que la astronomía en estas energías haya experimentado un desarrollo más lento en comparación con otras longitudes de onda. Hoy en día, el gran éxito de consecutivas misiones espaciales junto con el desarrollo y el perfeccionamiento de nuevas técnicas de detección en la superficie de la Tierra, ha supuesto excelentes resultados científicos y ha colocado a la astronomía de rayos gamma a un gran nivel, a

la par con el resto de campos de la astronomía.

Este trabajo está dedicado al estudio y la mejora del futuro Cherenkov Telescope Array (CTA), la próxima generación de detectores terrestres de rayos gamma, diseñados para observar los fotones con las energías más altas emitidas por fuentes cósmicas jamás observadas. Aquí se presentan estudios de sensibilidad realizados para la colaboración CTA, evaluados mediante el análisis detallado de simulaciones Monte Carlo a gran escala, calculando el rendimiento del observatorio junto con una estimación de su potencial en casos específicos dentro de la astrofísica. Junto con las pruebas y el desarrollo de las herramientas de análisis empleados, estos resultados son fundamentales para entender las capacidades que CTA tendrá en el futuro, la eficacia de los diferentes enfoques en la distribución de telescopios y el efecto sobre el rendimiento del lugar de construcción en relación con parámetros como la altitud o el campo geomagnético.

En cuanto a los diferentes diseños de CTA, se realizó una comparación imparcial de las diferentes disposiciones de telescopios propuestos por la colaboración, utilizando los catálogos de Fermi-LAT para pronosticar el comportamiento del observatorio sobre casos científicos específicos. Además, se estudió la aplicación de algoritmos de aprendizaje en el campo, comparando métodos alternativos para la reconstrucción de la energía y separación señal-ruido. También se han introducido nuevas aplicaciones de estos algoritmos, tales como la determinación de tipos de fuentes a través del reconocimiento de sus características espectrales.

El análisis presentado aquí tanto de CTA-N como de CTA-S representa el estudio más completo sobre las capacidades de CTA realizado por parte de la colaboración hasta la fecha. La experiencia adquirida con la mejora de este software guiará al futuro análisis de datos, comparando la sensibilidad alcanzada por las diferentes cadenas de análisis. A partir de estos resultados, los candidatos “2N” y “2Q” cumplen con los requisitos exigidos por la colaboración en sensibilidad, área efectiva y resolución angular y energética. Las simulaciones Monte Carlo proporcionan un banco de pruebas para los diferentes diseños propuestos por el consorcio CTA y los resultados presentados aquí demuestran que su correcta implementación alcanzaría el rendimiento deseado, así como su enorme potencial científico.

List of Figures

1.1	Cosmic-ray spectrum	7
1.2	Gamma-ray Production Mechanisms	9
1.3	Optical depth by photon-photon pair production as a function of the redshift	14
1.4	Pulsar γ -ray emission proposed regions.	17
1.5	AGN unified scheme	22
2.1	Cos-B Milky Way map in the γ -ray range.	29
2.2	Altitude of the atmospheric absorption along the energy bands of the electromagnetic spectrum	30
2.3	Extended Atmospheric Shower detection techniques	37
2.4	Differential sensitivity of current VHE gamma-ray detectors	38
2.5	Fermi-LAT full sky map after 5 years of operation	42
2.6	Number of detected sources for different astronomy energy bands vs time	46
2.7	Differential sensitivity for future VHE gamma-ray detectors	47
2.8	Angular resolution of present and future VHE gamma-ray detectors	48
2.9	The HAWC Observatory	50
2.10	Events simulated by HAWC.	51
3.1	Different IACT optics	61
3.2	Schematics of the stereoscopic imaging principle used by Imaging Atmospheric Cherenkov Telescope (IACT)s	63
3.3	Baseline design of the large-sized telescopes (LST)	65

LIST OF FIGURES

3.4	Baseline designs of the proposed MSTs	66
3.5	Baseline design of the 4m-small-sized telescopes (SST) types	68
3.6	Horizontal projection of 3 showers development: gamma, proton and iron	71
3.7	Arrival position of Cherenkov photons in the reflectors, reaching the camera plane	73
3.8	Pixel amplitude of two CTA telescope types	81
3.9	Effect on the direction reconstruction of different image cleaning levels	83
3.10	Schematics of Hillas parameterization.	85
3.11	Differential sensitivity in 50 hours for the “2Q” candidate layout . .	90
3.12	Simulated telescope positions of the first large-scale CTA MC pro- duction (Prod-1)	92
3.13	Telescope distribution of 4 proposed candidate arrays from the Prod- 1	94
3.14	Proposed layouts to compare CTA performance over different con- struction sites.	101
3.15	Proposed layouts for the Southern Hemisphere site, using 4m-Schwarzschild- Couder small-sized telescopes (SC-SST)s	101
3.16	Proposed mixed DC/Schwarzschild-Couder (SC)-medium-sized tele- scopes (MST)s layouts	103
3.17	Differential sensitivity of the “2Q” layout using alternative analysis chains.	104
3.18	CTA differential sensitivity over 50, 5 and 0.5 hours.	106
3.19	CTA angular and energy resolution.	107
3.20	CTA-S Telescope types contributions to the sensitivity.	108
3.21	CTA-S off-axis performance	109
3.22	CTA-N and CTA-S off-axis sensitivity ratios.	110
3.23	Angular resolution of the “2Q” candidate array	112
3.24	Pixel amplitude of two different MST types	114
3.25	Pixel amplitudes of the Schwarzschild-Couder medium-sized tele- scopes (SC-MST) for different image cleaning thresholds	115
3.26	Differential sensitivity and angular resolution of DC-MSTs vs SC- MSTs.	116

3.27	Differential sensitivity of “2KA” and “2KB” layouts	117
3.28	Differential sensitivity of mixed DC/SC-MST layouts	118
3.29	Angular resolution of mixed DC/SC-MST layouts	120
3.30	Off-axis sensitivity ratios of layouts with different MST types	122
3.31	Differential sensitivity of northern sites	125
3.32	Orthogonal GF maximum intensity along the zenith angle	126
3.33	Different steps used in the calculation of the CTA-N candidate sites averaged orthogonal GF effect	128
3.34	Ratio between the γ rates of different NSB levels after high level quality cuts are applied	131
3.35	Alternative layouts proposed for the CTA-N at the ORM site. . . .	133
3.36	Differential sensitivity of the standard CTA-N “2N” and the two proposed alternative layouts	134
3.37	Angular resolution of standard CTA-N “2N” and the two proposed alternative layouts	135
3.38	Standard Southern Hemisphere simulated telescope positions of the second large-scale CTA MC production (Prod-2)	136
3.39	Standard Northern Hemisphere simulated telescope positions of the second large-scale CTA MC production (Prod-2)	137
3.40	Simulated telescope positions of the 4m SSTs extension to the sec- ond large-scale CTA MC production (Prod-2)	138
4.1	Comparison of calculated significances between <i>CTA macros</i> and <i>GAEtools</i>	144
4.2	γ -ray spectra from the annihilation of 1 TeV WIMPs for different channels	147
4.3	Cumulative distributions of the number of <i>Fermi</i> -LAT galactic and extragalactic sources detectable by the CTA	149
4.4	Cumulative distribution of the number of <i>Fermi</i> -LAT extragalactic sources of known redshift detectable by the CTA	153
4.5	Number of <i>Fermi</i> -LAT extragalactic sources of known redshift de- tectable by the CTA for different site altitudes	155

LIST OF FIGURES

4.6	Observable AGNs from the 1FHL catalog that would be detected in 50 hours of observation by CTA-N and CTA-S as a function of the redshift	157
4.7	<i>Fermi</i> -LAT pulsars general profile with Prod-1 CTA sensitivity . . .	158
4.8	Simulated Crab pulsar SED within 50 h of observation with the CTA	159
4.9	Pulsars detectable by the future CTA in 50 hours of observation time assuming their spectra are extrapolated with the Crab pulsar power-law tail	160
4.10	Pulsars detectable by the Prod-2 “2A” layout in 50 hours assuming their spectra are extrapolated with a variable power-law tail index .	161
4.11	CTA sensitivities on the velocity-averaged annihilation cross-section as a function of the WIMP mass for different dSph galaxies	163
4.12	Minimum value of the astrophysical factor (J) required for a detection within 100 hours of observation as a function of the WIMP mass	164
4.13	Minimum boost factor required for a detection in 100 h and 250 h .	165
5.1	Relative energy resolution for different energy reconstruction methods	172
5.2	Reconstructed energy over estimated energy as a function of the estimated energy	173
5.3	Mean Gini decrease calculated by the RF of each different telescope type	174
5.4	ROC curves of the binary classifiers used for γ -hadron separation .	176
5.5	Q factors as a function of the energy for the different algorithms . .	177
5.6	γ -ray features ranked in order of importance	179
5.7	Properties of <i>Fermi</i> features	182
5.8	Distribution of outlyingness for all tested objects	184
5.9	Distributions of Powerlaw Index and Pivot Energy for identified/associated BL Lacs and FSRQs	189
5.10	Decision threshold P_{bzb} obtained with RF versus P_{bzb} estimated by SVM	192

5.11	Photon spectral index versus log flux above 100 MeV for identified/associated BL Lacs (dark contour) and FSRQs (light contour). Predicted BL Lacs (filled circles) and predicted FSRQs (open circles) from the AGU dataset are shown over the contours.	193
A.1	The Cherenkov effect	206
A.2	Development of a pure electromagnetic cascade, generated from an impinging primary γ -ray	210
A.3	γ -ray induced EAS longitudinal development	211
A.4	Cherenkov Angle and distance R_c of a perpendicularly impinging γ -ray along the atmosphere	212
A.5	Schematic figure of an hadronic shower	214
A.6	Different contributions to the Cosmic Ray spectrum between 100 MeV and 10 TeV collected	215
A.7	Location of generated secondaries of a primary particle of 10 GeV .	216
A.8	Geomagnetic effect on the distribution of Cherenkov photons at ground level	219
A.9	Lateral Distribution of Cherenkov photons at ground level for different geomagnetic field intensities	220

LIST OF FIGURES

List of Tables

2.1	Different γ -ray observational subranges	36
2.2	Specifications of the different generations of IACTs	43
3.1	Trace integration and image cleaning parameters for the CTA tele- scopes	82
3.2	Prod-1 telescopes parameters	93
3.3	Prod-1 candidate layouts	93
3.4	Prod-2 Configurations	97
3.5	Prod-2 simulated sites	99
3.6	Average orthogonal Geomagnetic field effect on the CTA-N candi- date sites	129
3.7	Overall performance of simulated sites of the Northern Hemisphere	129
4.1	Number of <i>Fermi</i> -LAT sources selected from the 2FGL catalog that would be detected by CTA	151
4.2	Number of detectable <i>Fermi</i> extragalactic sources with known red- shift for different array configurations (50 h of exposure time) . . .	152
4.3	Observable AGNs from the 1FHL that would be detected in 50h of observation by CTA-N and CTA-S.	156
5.1	Predictions and voting percentages for unassociated <i>Fermi</i> sources in the 2FGL, ordered by RA	183
5.2	Top outliers among high-latitude unassociated sources in the 2FGL	184
5.3	Predictions for <i>Fermi</i> AGN of uncertain type in the 2FGL, ordered by RA	191

LIST OF TABLES

5.4	Number of predicted AGU sources as a function of decision threshold.	192
5.5	Predictions for unassociated <i>Fermi</i> objects tagged as AGN by [203], ordered by RA.	195

References

- [1] T. C. Weekes, *Very high energy gamma-ray astronomy*, *Physics Reports* **160** (March, 1988) 0–0. 6
- [2] Perkins, D., *Particle Astrophysics*. Oxford University Press., 2011. 6
- [3] M. Longair, *High energy astrophysics*. Cambridge University Press, 2011. 6, 16
- [4] R. Millikan and C. G.H. *Phys. Rev.* **28** (1926) 851. 6, 28
- [5] V. Hess *Phys. Zeit.* **13** (1912) 1084. 6, 28
- [6] W. F. Hanlon, “Updated cosmic ray spectrum.”
<http://www.physics.utah.edu/~whanlon/spectrum.html>, 2015. 7
- [7] **Pierre Auger** Collaboration, J. Abraham *et. al.*, *Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects*, *Science* **318** (Nov., 2007) 938–, [arXiv:0711.2256]. 8
- [8] **Pierre Auger** Collaboration, J. Abraham *et. al.*, *Astrophysical Sources of Cosmic Rays and Related Measurements with the Pierre Auger Observatory*, *ArXiv e-prints* (June, 2009) [arXiv:0906.2347]. 8
- [9] I. Oya, *Observations of Active Galactic Nuclei with the MAGIC Telescope*. PhD dissertation, Universidad Complutense de Madrid, Atomic and Nuclear Physics Department, 2010. 9, 21, 211, 214

REFERENCES

- [10] D. Nieto, *Dark Matter Constrains from High Energy Astrophysical Observations*. PhD dissertation, Universidad Complutense de Madrid, Departamento de Física Atómica, Molecular y Nuclear, 2012. 13, 36, 147, 210
- [11] A. González, *Measurement of the gamma-ray opacity of the universe with the MAGIC telescopes*. PhD dissertation, Institut de Física d'Altes Energies, Universitat Autònoma de Barcelona, 2015. 13
- [12] A. Franceschini, G. Rodighiero, and M. Vaccari, *Extragalactic optical-infrared background radiation, its time evolution and the cosmic photon-photon opacity*, *Astronomy & Astrophysics* **487** (Sept., 2008) 837–852, [[arXiv:0805.1841](#)]. 13, 14, 141, 152
- [13] A. Domínguez *et. al.*, *Extragalactic background light inferred from AEGIS galaxy-SED-type fractions*, *Monthly Notices of the RAS* **410** (Feb., 2011) 2556–2578, [[arXiv:1007.1459](#)]. 13
- [14] F. W. Stecker and S. T. Scully, *A simple analytic treatment of the intergalactic absorption effect in blazar gamma-ray spectra*, *The Astrophysical Journal Letters* **652** (2006), no. 1 L9. 13
- [15] D. Mazin, *Constraints on Extragalactic Background Light from Cherenkov telescopes: status and perspectives for the next 5 years*, in *American Institute of Physics Conference Series* (D. Bastieri and R. Rando, eds.), vol. 1112 of *American Institute of Physics Conference Series*, pp. 111–120, Apr., 2009. [arXiv:0904.0774](#). 14, 156
- [16] E. Komatsu *et. al.*, *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophysical Journal, Supplement* **192** (Feb., 2011) 18, [[arXiv:1001.4538](#)]. 14
- [17] B. A. Reid, *Applying the halo model to large scale structure measurements of the luminous red galaxies: {SDSS} {DR7} preliminary results*, *Nuclear Physics B - Proceedings Supplements* **194** (2009), no. 0 129 – 132. New Horizons for Modern Cosmology: Proceedings of the Galileo Galilei Institute Conferences on Dark Matter and Dark Energy. 15

-
- [18] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, *The Aquarius Project: the subhaloes of galactic haloes*, *Monthly Notices of the RAS* **391** (Dec., 2008) 1685–1711, [arXiv:0809.0898]. 15, 25
- [19] B. Anderson, M. Kuhlen, J. Diemand, R. P. Johnson, and P. Madau, *Fermi-LAT Sensitivity to Dark Matter Annihilation in VIA Lactea II Substructure*, *Astrophysical Journal* **718** (Aug., 2010) 899–904, [arXiv:1006.1628]. 15
- [20] V. Springel, S. D. M. White, C. S. Frenk, J. F. Navarro, A. Jenkins, M. Vogelsberger, J. Wang, A. Ludlow, and A. Helmi, *Prospects for detecting supersymmetric dark matter in the Galactic halo*, *Nature* **456** (Nov., 2008) 73–76, [arXiv:0809.0894]. 15
- [21] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: evidence, candidates and constraints*, *Physics Reports* **405** (Jan., 2005) 279–390, [hep-ph/0404175]. 15
- [22] J. L. Feng, *Dark Matter Candidates from Particle Physics and Methods of Detection*, *Annual Review of Astronomy and Astrophysics* **48** (Sept., 2010) 495–545, [arXiv:1003.0904]. 15
- [23] M. Doro, J. Conrad, D. Emmanoulopoulos, M. A. Sánchez-Conde, J. A. Barrio, E. Birsin, J. Bolmont, P. Brun, S. Colafrancesco, S. H. Connell, J. L. Contreras, M. K. Daniel, M. Fornasa, M. Gaug, J. F. Glicenstein, A. González-Muñoz, T. Hassan, D. Horns, A. Jacholkowska, C. Jahn, R. Mazini, N. Mirabal, A. Moralejo, E. Moulin, D. Nieto, J. Ripken, H. Sandaker, U. Schwanke, G. Spengler, A. Stamerra, A. Viana, H.-S. Zechlin, S. Zimmer, and CTA Consortium, *Dark matter and fundamental physics with the Cherenkov Telescope Array*, *Astroparticle Physics* **43** (Mar., 2013) 189–214, [arXiv:1208.5356]. 15, 162, 165, 202
- [24] The Fermi-LAT Collaboration, *Fermi Large Area Telescope Third Source Catalog*, *ArXiv e-prints* (Jan., 2015) [arXiv:1501.0200]. 16, 17, 19, 40, 144

REFERENCES

- [25] E. Aliu *et. al.*, *Observation of Pulsed γ -Rays Above 25 GeV from the Crab Pulsar with MAGIC*, *Science* **322** (Nov., 2008) 1221–, [arXiv:0809.2998]. 17, 45, 145, 158
- [26] VERITAS Collaboration, *Detection of Pulsed Gamma Rays Above 100 GeV from the Crab Pulsar*, *Science* **334** (Oct., 2011) 69–, [arXiv:1108.3797]. 17, 145, 158, 159, 160
- [27] J. Aleksić *et. al.*, *Phase-resolved energy spectra of the Crab pulsar in the range of 50–400 GeV measured with the MAGIC telescopes*, *Astronomy & Astrophysics* **540** (Apr., 2012) A69, [arXiv:1109.6124]. 17, 140, 145, 159, 160
- [28] C. f. t. H. Stegmann, “H.E.S.S. Highlights.” <http://www.mpi-hd.mpg.de/hfm/HESS/201406-TeVPA-Amsterdam-nofilm.pdf>, 2014. 17
- [29] G. F. Krymskii, *A regular mechanism for the acceleration of charged particles on the front of a shock wave*, *Akademiia Nauk SSSR Doklady* **234** (June, 1977) 1306–1308. 18
- [30] M. Ackermann *et. al.*, *Detection of the Characteristic Pion-Decay Signature in Supernova Remnants*, *Science* **339** (Feb., 2013) 807–811, [arXiv:1302.3307]. 18, 41
- [31] D. Pooley, W. Lewin, S. Anderson, H. Baumgardt, A. Filippenko, *et. al.*, *Dynamical formation of close binary systems in globular clusters*, *Astrophysical Journal* **591** (2003) L131–L134, [astro-ph/0305003]. 18
- [32] H.E.S.S. Collaboration, A. Abramowski *et. al.*, *Very-high-energy gamma-ray emission from the direction of the Galactic globular cluster Terzan 5*, *ArXiv e-prints* (June, 2011) [arXiv:1106.4069]. 19
- [33] I. F. Mirabel and L. F. Rodríguez, *Microquasars in our Galaxy*, *Nature* **392** (Apr., 1998) 673–676. 19
- [34] I. F. Mirabel and L. F. Rodríguez, *A superluminal source in the Galaxy*, *Nature* **371** (Sept., 1994) 46–48. 19

-
- [35] Maraschi, L. and Treves, A., *A model for LS I 61° 303*, *MNRAS* **194** (1981) 1–5. 19
 - [36] **Fermi-LAT** Collaboration, A. A. Abdo *et. al.*, *Fermi LAT Observations of LS I +61deg303: First Detection of an Orbital Modulation in GeV Gamma Rays*, *ApJL* **701** (Aug., 2009) L123–L128, [[arXiv:0907.4307](#)]. 19
 - [37] **Fermi-LAT** Collaboration, A. A. Abdo *et. al.*, *Modulated High-Energy Gamma-Ray Emission from the Microquasar Cygnus X-3*, *Science* **326** (Dec., 2009) 1512–. 19
 - [38] **H.E.S.S.** Collaboration, Aharonian, F., and others, *Discovery of the Binary Pulsar PSR B1259-63 in Very-High-Energy Gamma Rays around Periastron with H.E.S.S.*, *A&A* **442** (2005) 1–10. 19
 - [39] **H.E.S.S.** Collaboration, Aharonian, F., and others, *Discovery of Very High Energy Gamma Rays Associated with an X-ray Binary*, *Science* **309** (2005) 746–749. 19
 - [40] **MAGIC** Collaboration, Albert, J., and others, *Variable Very-High-Energy Gamma-Ray Emission from the Microquasar LS I +61 303*, *Science* **312** (2006) 1771–1773. 19
 - [41] **MAGIC** Collaboration, Albert, J., and others, *Periodic very high energy gamma-ray emission from LS I +61 303 observed with the MAGIC telescope*, *ApJ* **693** (2009) 303–310. 19
 - [42] M. Ackermann and Fermi-LAT Collaboration, *Fermi establishes classical novae as a distinct class of gamma-ray sources*, *Science* **345** (Aug., 2014) 554–558, [[arXiv:1408.0735](#)]. 20
 - [43] L. Chomiuk *et. al.*, *Binary orbits as the driver of γ -ray emission and mass ejection in classical novae*, *Nature* **514** (Oct., 2014) 339–342, [[arXiv:1410.3473](#)]. 20
 - [44] K. N. Abazajian, *The consistency of Fermi-LAT observations of the galactic center with a millisecond pulsar population in the central stellar*

REFERENCES

- cluster*, *Journal of Cosmology and Astroparticle Physics* **3** (Mar., 2011) 10, [arXiv:1011.4275]. 20
- [45] C. Gordon and O. Macías, *Dark matter and pulsar model constraints from Galactic Center Fermi-LAT gamma-ray observations*, *Physical Review D* **88** (Oct., 2013) 083521, [arXiv:1306.5725]. 20
- [46] **Fermi-LAT** Collaboration, A. A. Abdo *et. al.*, *Fermi/Large Area Telescope Bright Gamma-Ray Source List*, *ApJS* **183** (July, 2009) 46–66, [arXiv:0902.1340]. 20
- [47] **HESS** Collaboration, Aharonian, F., and others, *Very high energy gamma rays from the direction of Sagittarius A**, *A&A* **425** (2004) L13–L17. 20
- [48] **MAGIC** Collaboration, J. Albert *et. al.*, *Observation of Gamma Rays from the Galactic Center with the MAGIC Telescope*, *ApJ* **638** (2006) L101–L104, [astro-ph/0512469]. 20
- [49] **Whipple** Collaboration, K. Kosack *et. al.*, *TeV Gamma-Ray Observations of the Galactic Center*, *ApJL* **608** (June, 2004) L97–L100, [astro-ph/0403422]. 20
- [50] Urry, C.M. and Padovani, P., *Unified Schemes for Radio Sources*, *Publications of the Astronomical Society of the Pacific* **107** (1995) 803. 21
- [51] J. Aleksić *et. al.*, *Black hole lightning due to particle acceleration at subhorizon scales*, *Science* **346** (Nov., 2014) 1080–1084, [arXiv:1412.4936]. 21, 45, 54
- [52] F. Aharonian *et. al.*, *A low level of extragalactic background light as revealed by γ -rays from blazars*, *Nature* **440** (Apr., 2006) 1018–1021, [astro-ph/0508073]. 21, 46
- [53] J. Aleksić *et. al.*, *MAGIC Observations and multiwavelength properties of the quasar 3C 279 in 2007 and 2009*, *Astronomy & Astrophysics* **530** (June, 2011) A4, [arXiv:1101.2522]. 21, 46

-
- [54] MAGIC Collaboration *et. al.*, *Very-High-Energy gamma rays from a Distant Quasar: How Transparent Is the Universe?*, *Science* **320** (June, 2008) 1752–, [arXiv:0807.2822]. 21, 46
- [55] V. Acciari *et. al.*, *Discovery of Very High Energy Gamma-ray Radiation from the BL Lac 1ES 0806+524*, *Astrophysical Journal Letters* **690** (Jan., 2009) L126–L129, [arXiv:0812.0978]. 21, 46
- [56] V. A. Acciari *et. al.*, *The Discovery of γ -Ray Emission from the Blazar RGB J0710+591*, *Astrophysical Journal Letters* **715** (May, 2010) L49–L55, [arXiv:1005.0041]. 21, 46
- [57] D. F. Torres and L. A. Anchordoqui, *Astrophysical origins of ultrahigh energy cosmic rays*, *Reports on Progress in Physics* **67** (Sept., 2004) 1663–1730, [astro-ph/0402371]. 22
- [58] **HESS** Collaboration, Acero, F., and others, *Detection of Gamma Rays from a Starburst Galaxy*, *Science* **326** (2009) 1080. 22
- [59] **VERITAS** Collaboration, V. A. Acciari *et. al.*, *A connection between star formation activity and cosmic rays in the starburst galaxy M82*, *Nature* **November 1** (2009). 22
- [60] **Fermi-LAT** Collaboration, A. A. Abdo *et. al.*, *Detection of Gamma-Ray Emission from the Starburst Galaxies M82 and NGC 253 with the Large Area Telescope on Fermi*, *ApJL* **709** (Feb., 2010) L152–L157, [arXiv:0911.5327]. 22
- [61] Klebesadel, R, Strong, I, and Olson, R, *Observations of gamma-ray bursts of cosmic origin*, *ApJ* **182** (1973) L85. 23
- [62] S. Campana *et. al.*, *The association of GRB 060218 with a supernova and the evolution of the shock wave*, *Nature* **442** (Aug., 2006) 1008–1010, [astro-ph/0603279]. 23
- [63] **Fermi-LAT** Collaboration, M. Ackermann *et. al.*, *Fermi Observations of High-energy Gamma-ray Emission from GRB 090217A*, *ApJL* **717** (July, 2010) L127–L132, [arXiv:1007.3409]. 23

REFERENCES

- [64] **MAGIC** Collaboration, A. Carosi *et. al.*, *MAGIC low energy observation of GRB090102 afterglow*, *ArXiv e-prints* (Oct., 2011) [[arXiv:1110.6848](#)]. 23
- [65] M. Garczarczyk, *First observations of the GRB prompt and early afterglow emission phase at the 100 GeV energy regime with the 17 m MAGIC Imaging Atmospheric Cherenkov Telescope*. PhD dissertation, Universitat Rostock, Physics Department Galileo Galilei, 2006. 23
- [66] A. Pinzke, C. Pfrommer, and L. Bergstrom, *Prospects of detecting gamma-ray emission from galaxy clusters: cosmic rays and dark matter annihilations*, *ArXiv e-prints* (May, 2011) [[arXiv:1105.3240](#)]. 24
- [67] M. Ackermann *et. al.*, *GeV Gamma-ray Flux Upper Limits from Clusters of Galaxies*, *ApJL* **717** (July, 2010) L71–L78, [[arXiv:1006.0748](#)]. 24
- [68] **MAGIC** Collaboration, J. Aleksić *et. al.*, *Constraining Cosmic Rays and Magnetic Fields in the Perseus Galaxy Cluster with TeV observations by the MAGIC telescopes*, *ArXiv e-prints* (Nov., 2011) [[arXiv:1111.5544](#)]. 24
- [69] F. Aharonian *et. al.*, *Very high energy gamma-ray observations of the galaxy clusters Abell 496 and Abell 85 with HESS*, *Astronomy & Astrophysics* **495** (Feb., 2009) 27–35, [[arXiv:0812.1638](#)]. 24, 162
- [70] V. A. Acciari *et. al.*, *VERITAS Upper Limit on the Very High Energy Emission from the Radio Galaxy NGC 1275*, *Astrophysical Journal Letters* **706** (Dec., 2009) L275–L280, [[arXiv:0911.0740](#)]. 24, 162
- [71] J. Aleksić *et. al.*, *Constraining cosmic rays and magnetic fields in the Perseus galaxy cluster with TeV observations by the MAGIC telescopes*, *Astronomy & Astrophysics* **541** (May, 2012) A99, [[arXiv:1111.5544](#)]. 24, 162
- [72] A. Abramowski and H. E. S. S. Collaboration, *Search for Dark Matter Annihilation Signals from the Fornax Galaxy Cluster with H.E.S.S.*, *Astrophysical Journal* **750** (May, 2012) 123, [[arXiv:1202.5494](#)]. 24, 162

-
- [73] Q. D. Wang, F. J. Lu, and E. V. Gotthelf, *G359.95-0.04: an energetic pulsar candidate near Sgr A**, *Monthly Notices of the RAS* **367** (Apr., 2006) 937–944, [astro-ph/0512643]. 25
- [74] A. A. Abdo and Fermi LAT Collaboration, *Observations of Milky Way Dwarf Spheroidal Galaxies with the Fermi-Large Area Telescope Detector and Constraints on Dark Matter Models*, *Astrophysical Journal* **712** (Mar., 2010) 147–158, [arXiv:1001.4531]. 25
- [75] F. Aharonian *et. al.*, *Observations of the Sagittarius dwarf galaxy by the HESS experiment and search for a dark matter signal*, *Astroparticle Physics* **29** (Feb., 2008) 55–62, [arXiv:0711.2369]. 25, 162
- [76] F. Aharonian *et. al.*, *A Search for a Dark Matter Annihilation Signal Toward the Canis Major Overdensity with H.E.S.S.*, *Astrophysical Journal* **691** (Jan., 2009) 175–181, [arXiv:0809.3894]. 25, 162
- [77] J. Aleksić *et. al.*, *Searches for dark matter annihilation signatures in the Segue 1 satellite galaxy with the MAGIC-I telescope*, *Journal of Cosmology and Astroparticle Physics* **6** (June, 2011) 35, [arXiv:1103.0477]. 25, 162
- [78] J. Aleksić *et. al.*, *Optimized dark matter searches in deep observations of Segue 1 with MAGIC*, *Journal of Cosmology and Astroparticle Physics* **2** (Feb., 2014) 8, [arXiv:1312.1535]. 25, 162
- [79] V. A. Acciari and VERITAS Collaboration, *VERITAS Search for VHE Gamma-ray Emission from Dwarf Spheroidal Galaxies*, *Astrophysical Journal* **720** (Sept., 2010) 1174–1180, [arXiv:1006.5955]. 25, 162
- [80] E. Aliu *et. al.*, *VERITAS deep observations of the dwarf spheroidal galaxy Segue 1*, *Physical Review D* **85** (Mar., 2012) 062001, [arXiv:1202.2144]. 25, 162
- [81] H.E.S.S. Collaboration, *H.E.S.S. constraints on dark matter annihilations towards the sculptor and carina dwarf galaxies*, *Astroparticle Physics* **34** (Mar., 2011) 608–616, [arXiv:1012.5602]. 25, 46

REFERENCES

- [82] T. Wulf, *The thread electrometers*, . 28
- [83] D. Pacini *Rend. Acc. Lincei* **18** (1909) 123. 28
- [84] J. Dreitlein and H. Primakoff, *Radiative Decay of the Neutral K Meson: $K^0 \rightarrow \gamma - \gamma$* , *Phys.Rev.* **124** (1961) 268–273. 29
- [85] S. Hayakawa, *Polarization of Cosmic-Ray mu Mesons: Theory*, *Phys.Rev.* **108** (1957) 1533–1537. 29
- [86] P. Morrison and L. I. Schiff, *Radiative K Capture*, *Physical Review* **58** (July, 1940) 24–26. 29
- [87] D. J. Thompson *et. al.*, *Final SAS-2 gamma-ray results on sources in the galactic anticenter region*, *Astrophysical Journal* **213** (Apr., 1977) 252–262. 29
- [88] J. B. G. M. Bloemen *et. al.*, *The radial distribution of galactic gamma rays. II - The distribution of cosmic-ray electrons and nuclei in the outer galaxy*, *Astronomy & Astrophysics* **135** (June, 1984) 12–22. 29
- [89] B. N. Swanenburg *et. al.*, *Second COS B catalog of high-energy gamma-ray sources*, *Astrophysical Journal Letters* **243** (Jan., 1981) L69–L73. 29
- [90] C. L. X. J. de la Vallée Poussin, 1. A strong form of the prime number theorem, 19th century. 31
- [91] **VERITAS** Collaboration, J. Kildea *et. al.*, *The whipple observatory 10m gamma-ray telescope, 1997 - 2006*, *Astroparticle Physics* **28** (2007), no. 2 182 – 195. 33
- [92] **MAGIC** Collaboration, “The MAGIC Telescopes Homepage.” <http://www.magic.mppmu.mpg.de/>, 2011. 33
- [93] **H.E.S.S.** Collaboration, “H.E.S.S. Homepage.” <http://www.mpi-hd.mpg.de/hfm/HESS/>, 2011. 33
- [94] **VERITAS** Collaboration, “VERITAS Homepage.” <http://veritas.sao.arizona.edu/>, 2011. 33, 38, 48

-
- [95] **MILAGRO** Collaboration, “The Milagro Gamma-Ray Observatory Homepage.” <http://www.lanl.gov/milagro/>, 2011. 33
- [96] **HAWC** Collaboration, “HAWC Observatory Homepage.” <http://www.hawc-observatory.org/>, 2014. 33, 47, 48, 50, 51
- [97] **Tibet-AS** Collaboration, “Tibet-AS Gamma Experiment Homepage.” <http://www.icrr.u-tokyo.ac.jp/em/>, 2014. 33
- [98] **Argo-YBJ** Collaboration, “Argo-YBJ Homepage.” <http://argo.na.infn.it/>, 2014. 33
- [99] **HiRes** Collaboration, “High Resolution Fly’s Eye Homepage.” <http://www.cosmic-ray.org/>, 2014. 34
- [100] **Pierre Auger** Collaboration, “The Pierre Auger Observatory Homepage.” <http://www.auger.org/>, 2014. 34
- [101] M. Ackermann *et. al.*, *The Fermi Large Area Telescope on Orbit: Event Classification, Instrument Response Functions, and Calibration*, *Astrophysical Journal Supplement* **203** (Nov., 2012) 4, [arXiv:1206.1896]. 38, 47, 48
- [102] J. Aleksić *et. al.*, *Performance of the MAGIC stereo system obtained with Crab Nebula data*, *Astroparticle Physics* **35** (Feb., 2012) 435–448, [arXiv:1108.1477]. 38, 48, 69
- [103] **Fermi** Collaboration, “Fermi Gamma-ray Space Telescope Homepage.” <http://fermi.gsfc.nasa.gov/>, 2014. 37
- [104] **Fermi-LAT** Collaboration, “The Fermi Large Area Telescope Homepage.” <http://www-glast.stanford.edu/>, 2014. 39
- [105] A. A. Abdo *et. al.*, *The Fermi Gamma-Ray Space Telescope Discovers the Pulsar in the Young Galactic Supernova Remnant CTA 1*, *Science* **322** (Nov., 2008) 1218–, [arXiv:0810.3562]. 41

REFERENCES

- [106] Fermi-LAT collaboration and Fermi-GBM collaboration, *Fermi-LAT Observations of the Gamma-ray Burst GRB 130427A*, *ArXiv e-prints* (Nov., 2013) [[arXiv:1311.5623](#)]. 41
- [107] M. Su, T. R. Slatyer, and D. P. Finkbeiner, *Giant Gamma-ray Bubbles from Fermi-LAT: Active Galactic Nucleus Activity or Bipolar Galactic Wind?*, *ApJ* **724** (Dec., 2010) 1044–1082, [[arXiv:1005.5480](#)]. 41
- [108] **Fermi** Collaboration, “Fermi’s Five-year View of the Gamma-ray Sky.” <http://svs.gsfc.nasa.gov/vis/a010000/a011300/a011342/>, 2014. 42
- [109] **Fermi-LAT** Collaboration, The Fermi-LAT Collaboration, *Fermi Large Area Telescope Second Source Catalog*, *ArXiv e-prints* (Aug., 2011) [[arXiv:1108.1435](#)]. 42, 179, 188, 189
- [110] T. C. Weekes *et. al.*, *Observation of TeV gamma rays from the Crab nebula using the atmospheric Cerenkov imaging technique*, *Astrophysical Journal* **342** (July, 1989) 379–395. 43
- [111] **HEGRA** Collaboration, A. Kohnle *et. al.*, *Stereoscopic imaging of air showers with the first two HEGRA Cherenkov telescopes*, *Astroparticle Physics* **5** (Aug., 1996) 119–131. 43
- [112] J. Hinton, *Ground-based gamma-ray astronomy with Cherenkov telescopes*, *New Journal of Physics* **11** (May, 2009) 055005–+, [[arXiv:0803.1609](#)]. 43
- [113] Nature, “The world’s top ten telescopes revealed.” <http://www.nature.com/news/2009/090206/full/news.2009.81.html>, 2009. 45
- [114] S. Funk, *The status of gamma-ray astronomy*, *ArXiv e-prints* (Apr., 2012) [[arXiv:1204.4529](#)]. 46
- [115] F. Aharonian *et. al.*, *The H.E.S.S. Survey of the Inner Galaxy in Very High Energy Gamma Rays*, *Astrophysical Journal* **636** (Jan., 2006) 777–797, [[astro-ph/0510397](#)]. 45

-
- [116] F. Aharonian *et. al.*, *A New Population of Very High Energy Gamma-Ray Sources in the Milky Way*, *Science* **307** (Mar., 2005) 1938–1942, [astro-ph/0504380]. 45
- [117] VERITAS Collaboration *et. al.*, *Detection of Pulsed Gamma Rays Above 100 GeV from the Crab Pulsar*, *Science* **334** (Oct., 2011) 69–, [arXiv:1108.3797]. 45
- [118] A. Abramowski *et. al.*, *Search for a Dark Matter Annihilation Signal from the Galactic Center Halo with H.E.S.S.*, *Physical Review Letters* **106** (Apr., 2011) 161301, [arXiv:1103.3266]. 46
- [119] J. Aleksić *et. al.*, *Optimized dark matter searches in deep observations of Segue 1 with MAGIC*, *Journal of Cosmology and Astroparticle Physics* **2** (Feb., 2014) 8, [arXiv:1312.1535]. 46
- [120] V. A. Acciari *et. al.*, *VERITAS Search for VHE Gamma-ray Emission from Dwarf Spheroidal Galaxies*, *Astrophysical Journal* **720** (Sept., 2010) 1174–1180, [arXiv:1006.5955]. 46
- [121] F. Aharonian *et. al.* arXiv:0811.3894. 46
- [122] S. Funk, J. A. Hinton, and CTA Consortium, *Comparison of Fermi-LAT and CTA in the region between 10-100 GeV*, *Astroparticle Physics* **43** (Mar., 2013) 348–355, [arXiv:1205.0832]. 47, 48
- [123] R. Zanin, *Observation of the Crab pulsar wind nebula and microquasar candidates with MAGIC*. PhD dissertation, Universitat Autònoma de Barcelona, Institut de Física d’Altes Energies, 2011. 58
- [124] V. V. Vassiliev and S. J. Fegan, *Schwarzschild-Couder two-mirror telescope for ground-based γ -ray astronomy*, *International Cosmic Ray Conference* **3** (2008) 1445–1448, [arXiv:0708.2741]. 61
- [125] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, *CORSIKA: a Monte Carlo code to simulate extensive air showers*. Forschungszentrum Karlsruhe GmbH, Feb., 1998. 63, 70

REFERENCES

- [126] F. Schmidt, “CORSIKA Shower Images.”
<http://www.ast.leeds.ac.uk/~fs/showerimages.html>, 2014. 63, 71
- [127] A. M. Hillas, *Cerenkov light images of EAS produced by primary gamma*, *International Cosmic Ray Conference* **3** (Aug., 1985) 445–448. 62, 76
- [128] C.-C. Lu and for the H. E. S. S. Collaboration, *Improving the H.E.S.S. angular resolution using the Disp method*, *ArXiv e-prints* (Oct., 2013) [arXiv:1310.1200]. 69
- [129] V. B. Anykeyev *et. al.*, *Comparative investigation of unfolding methods*, *Nuclear Instruments and Methods in Physics Research A* **303** (June, 1991) 350–369. 69
- [130] J. Albert *et. al.* 69
- [131] A. Moralejo, M. Gaug, E. Carmona, P. Colin, C. Delgado, S. Lombardi, D. Mazin, V. Scalzotto, J. Sitarek, D. Tescaro, and for the MAGIC collaboration, *MARS, the MAGIC Analysis and Reconstruction Software*, *ArXiv e-prints* (July, 2009) [arXiv:0907.0943]. 69
- [132] J. N. Capdevielle *et. al.*, *The Karlsruhe extensive air shower simulation code CORSIKA*. Nov., 1992. 70
- [133] T. Antoni and (KASCADE-Collaboration), *The cosmic-ray experiment KASCADE*, *Nuclear Instruments and Methods in Physics Research A* **513** (Nov., 2003) 490–510. 70
- [134] J. Shiers, *The Worldwide LHC Computing Grid (worldwide LCG)*, *Computer Physics Communications* **177** (July, 2007) 219–223. 71
- [135] K. Bernlöhr, *Simulation of imaging atmospheric Cherenkov telescopes with CORSIKA and sim_telarray*, *Astroparticle Physics* **30** (Oct., 2008) 149–158, [arXiv:0808.2253]. 71, 80
- [136] T.-P. Li and Y.-Q. Ma, *Analysis methods for results in gamma-ray astronomy*, *Astrophysical Journal* **272** (1983) 317–324. 75, 142

-
- [137] A. Daum, G. Hermann, M. Heß, and et al., *First results on the performance of the HEGRA IACT array*, *Astroparticle Physics* **8** (Dec., 1997) 1–11. 77
- [138] W. Hofmann, I. Jung, A. Konopelko, H. Krawczynski, H. Lampeitl, and G. Pühlhofer, *Comparison of techniques to reconstruct VHE gamma-ray showers from multiple stereoscopic Cherenkov images*, *Astroparticle Physics* **12** (Nov., 1999) 135–143, [astro-ph/9904234]. 77
- [139] F. A. Aharonian, W. Hofmann, A. K. Konopelko, and H. J. Völk, *The potential of ground based arrays of imaging atmospheric Cherenkov telescopes. I. Determination of shower parameters*, *Astroparticle Physics* **6** (Mar., 1997) 343–368. 78
- [140] M. Wood, T. Jogler, J. Dumm, and S. Funk, *Monte carlo studies of medium-size telescope designs for the cherenkov telescope array*, *Astroparticle Physics* (Nov, 2014). 78, 82
- [141] Bretz, T. and Wagner, R.M., *The MAGIC Analysis and Reconstruction Software*, *Proceedings of the 28th International Cosmic Ray Conference, Tsukuba, japan* **5** (2003) 2947–2950. 79
- [142] **MAGIC** Collaboration, S. Lombardi, K. Berger, P. Colin, A. Diago Ortega, S. Klepser, *et. al.*, *Advanced stereoscopic gamma-ray shower analysis with the MAGIC telescopes*, *ArXiv e-prints* (Sept., 2011) [arXiv:1109.6195]. 79
- [143] R. Brun and F. Rademakers, *ROOT - An object oriented data analysis framework*, *Nuclear Instruments and Methods in Physics Research A* **389** (Feb., 1997) 81–86. 79
- [144] M. Rissi, *Detection of Pulsed Very High Energy Gamma-Rays from the Crab Pulsar with the MAGIC telescope using an Analog Sum Trigger*. PhD dissertation, ETH Zurich, 2009. 82, 115
- [145] F. A. Aharonian *et. al.*, *The Energy Spectrum of TEV Gamma Rays from the Crab Nebula as Measured by the HEGRA System of Imaging Air*

REFERENCES

- Cerenkov Telescopes*, *Astrophysical Journal* **539** (Aug., 2000) 317–324, [astro-ph/0003182]. 88
- [146] T. Sanuki *et. al.*, *Precise Measurement of Cosmic-Ray Proton and Helium Spectra with the BESS Spectrometer*, *Astrophysical Journal* **545** (Dec., 2000) 1135–1142, [astro-ph/0002481]. 88
- [147] K. Bernlöhner, *Impact of atmospheric parameters on the atmospheric Cherenkov technique**, *Astroparticle Physics* **12** (Jan., 2000) 255–268, [astro-ph/9908093]. 97
- [148] F. X. Kneizys *et. al.*, *The MODTRAN 2/3 report and LOWTRAN 7 model*. Phillips Laboratory, Hanscom AFB, MA 01731, U.S.A., 1996, . 97
- [149] S. C. Commichau, A. Biland, D. Kranich, and *et al.*, *Geomagnetic Field Effects on the Imaging Air Shower Cherenkov Technique*, *International Cosmic Ray Conference* **3** (2008) 1357–1360, [arXiv:0709.1251]. 126
- [150] NASA, “A General Package of Software to Manipulate FITS Files.” <http://heasarc.gsfc.nasa.gov/fertools/>, 1995. 140
- [151] M. Ackermann *et. al.*, *The First Fermi-LAT Catalog of Sources above 10 GeV*, *Astrophysical Journal, Supplement* **209** (Dec., 2013) 34, [arXiv:1306.6772]. 146
- [152] J. A. R. Cembranos, A. de La Cruz-Dombriz, A. Dobado, R. A. Lineros, and A. L. Maroto, *Photon spectra from WIMP annihilation*, *Physical Review D* **83** (Apr., 2011) 083507, [arXiv:1009.4936]. 146, 147, 162
- [153] J. Aleksić, J. Rico, and M. Martinez, *Optimized analysis method for indirect dark matter searches with imaging air Cherenkov telescopes*, *Journal of Cosmology and Astroparticle Physics* **10** (Oct., 2012) 32, [arXiv:1209.5589]. 147
- [154] W. A. Rolke, A. M. López, and J. Conrad, *Limits and confidence intervals in the presence of nuisance parameters*, *Nuclear Instruments and Methods in Physics Research A* **551** (Oct., 2005) 493–503, [physics/0403059]. 147

-
- [155] M. Actis and et al., *Design concepts for the Cherenkov Telescope Array CTA: an advanced facility for ground-based high-energy gamma-ray astronomy*, *Experimental Astronomy* **32** (Dec., 2011) 193–316, [arXiv:1008.3703]. 148
- [156] P. L. Nolan, A. A. Abdo, et. al., *Fermi Large Area Telescope Second Source Catalog*, *Astrophysical Journal* **199** (Apr., 2012) 31, [arXiv:1108.1435]. 148
- [157] A. A. Abdo, M. Ajello, A. Allafort, L. Baldini, J. Ballet, G. Barbiellini, M. G. Baring, D. Bastieri, A. Belfiore, R. Bellazzini, and et al., *The Second Fermi Large Area Telescope Catalog of Gamma-Ray Pulsars*, *Astrophysical Journal, Supplement* **208** (Oct., 2013) 17, [arXiv:1305.4385]. 158
- [158] J. Aleksi et. al., *Observations of the Crab Pulsar between 25 and 100 GeV with the MAGIC I Telescope*, *The Astrophysical Journal* **742** (2011), no. 1 43. 161
- [159] A. Abramowski et. al., *Search for Photon-Linelike Signatures from Dark Matter Annihilations with H.E.S.S.*, *Physical Review Letters* **110** (Jan., 2013) 041301, [arXiv:1301.1173]. 162
- [160] M. G. Walker, M. Mateo, E. W. Olszewski, J. Peñarrubia, N. Wyn Evans, and G. Gilmore, *A Universal Mass Profile for Dwarf Spheroidal Galaxies?*, *Astrophysical Journal* **704** (Oct., 2009) 1274–1287, [arXiv:0906.0341]. 162
- [161] J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, and et al., *The Sequence of the Human Genome*, *Science* **291** (Feb., 2001) 1304–1351. 167
- [162] The C. Elegans Sequencing Consortium, *Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology*, *Science* **282** (Dec., 1998) 2012. 167
- [163] A. Gofeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W.

REFERENCES

- Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver, *Life with 6000 Genes*, *Science* **274** (Oct., 1996) 546. 167
- [164] H. N. Elomaa T, *An experimental comparison of inducing decision trees and decision lists in noisy domains*, *Proceedings of the 4th European Working Session on Learning* (Dec., 1989) 59–69. 168
- [165] I. Kononenko, *Machine learning for medical diagnosis: History, state of the art and perspective*, *Artif. Intell. Med.* **23** (Aug., 2001) 89–109. 168
- [166] G. Aad *et. al.*, *Observation of a new particle in the search for the standard model higgs boson with the {ATLAS} detector at the {LHC}*, *Physics Letters B* **716** (2012), no. 1 1 – 29. 168
- [167] S. Chatrchyan *et. al.*, *Observation of a new boson at a mass of 125 gev with the {CMS} experiment at the {LHC}*, *Physics Letters B* **716** (2012), no. 1 30 – 61. 168
- [168] S. Ohm, C. van Eldik, and K. Egberts, *γ /hadron separation in very-high-energy γ -ray astronomy using a multivariate analysis method*, *Astroparticle Physics* **31** (June, 2009) 383–391, [[arXiv:0904.1136](#)]. 168
- [169] J. Albert *et. al.*, *Implementation of the Random Forest method for the Imaging Atmospheric Cherenkov Telescope MAGIC*, *Nuclear Instruments and Methods in Physics Research A* **588** (Apr., 2008) 424–432, [[arXiv:0709.3719](#)]. 168
- [170] L. Breiman, *Random forests*, *Machine Learning* **45** (2001) 5–32. 10.1023/A:1010933404324. 169, 170, 180, 183, 188, 194
- [171] A. Liaw and M. Wiener, *Classification and regression by randomforest*, *R News* **2** (2002), no. 3 18–22. 169, 183, 188, 189, 194
- [172] J. H. Friedman, *Stochastic gradient boosting*, *Comput. Stat. Data Anal.* **38** (Feb., 2002) 367–378. 169

-
- [173] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, *PoS ACAT* (2007) 040, [[physics/0703039](#)]. 169
- [174] V. V. N., *Statistics for Engineering and Information Science*. Springer, 1999. 170
- [175] C. Cortes and V. N. Vapnik, *Support-Vector Networks*. 1995. 170
- [176] C.-C. Chang and C.-J. Lin, *Libsvm: A library for support vector machines*, *ACM Trans. Intell. Syst. Technol.* **2** (May, 2011) 27:1–27:27. 170
- [177] J. Diemand, M. Kuhlen, P. Madau, M. Zemp, B. Moore, D. Potter, and J. Stadel, *Clumps and streams in the local dark matter distribution*, *Nature* **454** (Aug., 2008) 735–738, [[arXiv:0805.1244](#)]. 178
- [178] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, *The Aquarius Project: the subhaloes of galactic haloes*, *Monthly Notices of the RAS* **391** (Dec., 2008) 1685–1711, [[arXiv:0809.0898](#)]. 178
- [179] A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, *Where Are the Missing Galactic Satellites?*, *Astrophysical Journal* **522** (Sept., 1999) 82–92, [[astro-ph/9901240](#)]. 178
- [180] B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, and P. Tozzi, *Dark Matter Substructure within Galactic Halos*, *Astrophysical Journal Letters* **524** (Oct., 1999) L19–L22, [[astro-ph/9907411](#)]. 178
- [181] M. Kuhlen, P. Madau, and J. Silk, *Exploring Dark Matter with Milky Way Substructure*, *Science* **325** (Aug., 2009) 970–, [[arXiv:0907.0005](#)]. 178
- [182] D. Nieto, V. Martínez, N. Mirabal, J. A. Barrio, K. Satalecka, S. Pardo, and I. Lozano, *A search for possible dark matter subhalos as IACT targets in the First Fermi-LAT Source Catalog*, *ArXiv e-prints* (Oct., 2011) [[arXiv:1110.4744](#)]. 178

REFERENCES

- [183] A. V. Belikov, M. R. Buckley, and D. Hooper, *Searching for dark matter subhalos in the Fermi-LAT second source catalog*, *Physical Review D* **86** (Aug., 2012) 043504, [[arXiv:1111.2613](#)]. 178
- [184] H.-S. Zechlin and D. Horns, *Unidentified sources in the Fermi-LAT second source catalog: the case for DM subhalos*, *Journal of Cosmology and Astroparticle Physics* **11** (Nov., 2012) 50, [[arXiv:1210.3852](#)]. 178, 182
- [185] M. Ackermann *et. al.*, *Search for Dark Matter Satellites Using Fermi-LAT*, *Astrophysical Journal* **747** (Mar., 2012) 121, [[arXiv:1201.2691](#)]. 178
- [186] L. Bergström, T. Bringmann, M. Eriksson, and M. Gustafsson, *Gamma Rays from Heavy Neutralino Dark Matter*, *Physical Review Letters* **95** (Dec., 2005) 241301, [[hep-ph/0507229](#)]. 178
- [187] M. Ackermann *et. al.*, *A Statistical Approach to Recognizing Source Classes for Unassociated Sources in the First Fermi-LAT Catalog*, *Astrophysical Journal* **753** (July, 2012) 83, [[arXiv:1108.1202](#)]. 178, 179, 180, 181, 187
- [188] N. Mirabal, D. Nieto, and S. Pardo, *The exotic fraction among unassociated Fermi sources*, *ArXiv e-prints* (July, 2010) [[arXiv:1007.2644](#)]. 179
- [189] M. Ackermann *et. al.*, *The Second Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope*, *Astrophysical Journal* **743** (Dec., 2011) 171, [[arXiv:1108.1420](#)]. 179, 186, 188
- [190] V. Svetnik *et. al.*, *Random forest: A classification and regression tool for compound classification and qsar modeling*, *Journal of Chemical Information and Computer Sciences* **43** (2003), no. 6 1947–1958, [<http://dx.doi.org/10.1021/ci034160g>]. PMID: 14632445. 180
- [191] C. Ling, C. X. Ling, and C. Li, *Data mining for direct marketing: Problems and solutions*, in *In Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98*, pp. 73–79, AAAI Press, 1998. 180

-
- [192] N. Chawla, *Data mining for imbalanced datasets: An overview*, in *Data Mining and Knowledge Discovery Handbook* (O. Maimon and L. Rokach, eds.), pp. 853–867. Springer US, 2005. 180
- [193] E. A. Baltz, J. E. Taylor, and L. L. Wai, *Can Astrophysical Gamma-Ray Sources Mimic Dark Matter Annihilation in Galactic Satellites?*, *Astrophysical Journal Letters* **659** (Apr., 2007) L125–L128, [astro-ph/0610731]. 182
- [194] D. Hooper and T. Linden, *Gamma rays from the galactic center and the wmap haze*, *Phys. Rev. D* **83** (Apr, 2011) 083517. 185
- [195] M. Kerr *et. al.*, *Five New Millisecond Pulsars from a Radio Survey of 14 Unidentified Fermi-LAT Gamma-Ray Sources*, *Astrophysical Journal Letters* **748** (Mar., 2012) L2, [arXiv:1201.5160]. 185
- [196] S. L. A. C. C. Carroll R. J., Ruppert D., *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman and Hall, 2006. 185
- [197] A. N. Morgan, J. Long, J. W. Richards, T. Broderick, N. R. Butler, and J. S. Bloom, *Rapid, Machine-learned Resource Allocation: Application to High-redshift Gamma-Ray Burst Follow-up*, *Astrophysical Journal* **746** (Feb., 2012) 170, [arXiv:1112.3654]. 185
- [198] M. S. Shaw, R. W. Romani, S. E. Healey, G. Cotter, P. F. Michelson, and A. C. S. Readhead, *Optical Spectroscopy of Bright Fermi LAT Blazars*, *Astrophysical Journal* **704** (Oct., 2009) 477–484, [arXiv:0908.2996]. 186
- [199] B. Sbarufatti, A. Treves, and R. Falomo, *Imaging Redshifts of BL Lacertae Objects*, *Astrophysical Journal* **635** (Dec., 2005) 173–179, [astro-ph/0508200]. 186
- [200] A. M. Meisner and R. W. Romani, *Imaging Redshift Estimates for BL Lacertae Objects*, *Astrophysical Journal* **712** (Mar., 2010) 14–25, [arXiv:1002.1343]. 186

REFERENCES

- [201] A. Rau *et. al.*, *BL Lacertae objects beyond redshift 1.3 - UV-to-NIR photometry and photometric redshift for Fermi/LAT blazars*, *Astronomy & Astrophysics* **538** (Feb., 2012) A26, [[arXiv:1112.0025](#)]. 186
- [202] F. Massaro, R. D’Abrusco, G. Tosti, M. Ajello, D. Gasparrini, J. E. Grindlay, and H. A. Smith, *The WISE Gamma-Ray Strip Parameterization: The Nature of the Gamma-Ray Active Galactic Nuclei of Uncertain Type*, *Astrophysical Journal* **750** (May, 2012) 138, [[arXiv:1203.1330](#)]. 187, 196, 197
- [203] N. Mirabal, V. Frías-Martinez, T. Hassan, and E. Frías-Martinez, *Fermi’s SIBYL: mining the gamma-ray sky for dark matter subhaloes*, *Monthly Notices of the RAS* **424** (July, 2012) L64–L68, [[arXiv:1205.4825](#)]. 187, 188, 195, 203, 232
- [204] K. J. Lee, L. Guillemot, Y. L. Yue, M. Kramer, and D. J. Champion, *Application of the Gaussian mixture model in pulsar astronomy - pulsar classification and candidates ranking for the Fermi 2FGL catalogue*, *Monthly Notices of the RAS* **424** (Aug., 2012) 2832–2840, [[arXiv:1205.6221](#)]. 187
- [205] A. A. Abdo, M. Ackermann, I. Agudo, M. Ajello, H. D. Aller, M. F. Aller, E. Angelakis, A. A. Arkharov, M. Axelsson, U. Bach, and *et al.*, *The Spectral Energy Distribution of Fermi Bright Blazars*, *Astrophysical Journal* **716** (June, 2010) 30–70, [[arXiv:0912.2040](#)]. 188
- [206] G. Dubus, J. L. Contreras, S. Funk, Y. Gallant, T. Hassan, J. Hinton, Y. Inoue, J. Knödlseider, P. Martin, N. Mirabal, M. de Naurois, M. Renaud, and CTA Consortium, *Surveys with the Cherenkov Telescope Array*, *Astroparticle Physics* **43** (Mar., 2013) 317–330, [[arXiv:1208.5686](#)]. 197, 201
- [207] H. Sol, A. Zech, C. Boisson, U. Barres de Almeida, J. Biteau, J.-L. Contreras, B. Giebels, T. Hassan, Y. Inoue, K. Katarzyński, H. Krawczynski, N. Mirabal, J. Poutanen, F. Rieger, T. Totani, W. Benbow, M. Cerruti, M. Errando, L. Fallon, E. de Gouveia Dal Pino,

- J. A. Hinton, S. Inoue, J.-P. Lenain, A. Neronov, K. Takahashi, H. Takami, R. White, and CTA Consortium, *Active Galactic Nuclei under the scrutiny of CTA*, *Astroparticle Physics* **43** (Mar., 2013) 215–240, [arXiv:1304.3024]. 201
- [208] E. de Oña-Wilhelmi, B. Rudak, J. A. Barrio, J. L. Contreras, Y. Gallant, D. Hadasch, T. Hassan, M. Lopez, D. Mazin, N. Mirabal, G. Pedalletti, M. Renaud, R. de los Reyes, D. F. Torres, and CTA Consortium, *Prospects for observations of pulsars and pulsar wind nebulae with CTA*, *Astroparticle Physics* **43** (Mar., 2013) 287–300, [arXiv:1209.0357]. 202
- [209] T. Hassan, N. Mirabal, J. L. Contreras, and I. Oya, *Gamma-ray active galactic nucleus type through machine-learning algorithms*, *Monthly Notices of the RAS* **428** (Jan., 2013) 220–225, [arXiv:1209.4359]. 203
- [210] A. Moralejo, *Búsqueda de fuentes cósmicas de radiación gamma de muy alta energía con el detector AIROBICC*. PhD dissertation, Universidad Complutense de Madrid, Departamento de Física Atómica, Molecular y Nuclear, 2000. 205
- [211] M. Lopez, *Astronomía Gamma con el Telescopio MAGIC: Observaciones de la Nebulosa y Pulsar del Cangrejo*. PhD dissertation, Universidad Complutense de Madrid, Atomic and Nuclear Physics Department, 2006. 206, 215
- [212] J. V. Jelley, *Cherenkov radiation and its applications*, Pergamon Press. 208
- [213] W. Heitler, *Quantum Theory of Radiation*, Dover Press (Mar., 1954). 209
- [214] Rossi, B. and Greisen, K., *Cosmic Ray Theory*, *Reviews of Modern Physics* **13** (1941) 240–309. 210
- [215] T. K. Gaisser, *Cosmic Rays and Particle Physics*. Cambridge University Press. 211
- [216] B. Wiebel-Sooth, *Chemical composition in high energy Cosmic Rays*. PhD dissertation, University of Wuppertal, 1998. 215

REFERENCES

- [217] V. Sahakian, F. Aharonian, and A. Akhperjanian. 216, 217
- [218] G. Conconić *Phys. Rev* **95** (1953) 1705. 217
- [219] R. d. l. Reyes, *Search for Gamma-ray emission from pulsars with the MAGIC telescope: Sensitivity studies, data check and data analysis*. PhD dissertation, Universidad Complutense de Madrid, Atomic and Nuclear Physics Department, 2009. 219, 220